UNIVERSAL
LIBRARY

OU_156005

UNIVERSAL
LIBRARY

Osmania University Library.

Hydari Collection

Accession No. G3391

Call No. 520.1

LECTURES ON QUANTUM MECHANICS

LECTURES ON
QUANTUM MECHANICS

VOLUME I

BY

M. RAZIUDDIN SIDDIQI
M.A. (Cantab.), Ph.D. (Leipzig.)
Fellow of the National Institute of Sciences, India
Professor of Mathematics, Osmania University

OSMANIA UNIVERSITY
HYDERABAD-DECCAN
1938

All Rights Reserved

PRINTED AT
THE BANGALORE PRESS, BANGALORE CITY
1938

DEDICATED

TO

PROFESSOR WERNER HEISENBERG

PREFACE

THIS book is based on a series of lectures given at the Osmania University, and is meant for those students who are taking up the subject for the first time. While there is no lack of good books either on the popular or on the advanced side, there is ample scope for books giving a mathematical introduction to the beginner, presenting the fundamental principles in the light of their historical development and working out the mathematical theory as fully as possible.

A knowledge of only the elements of classical dynamics and electrodynamics as well as of vector analysis, is assumed on the part of the reader. In order to make the book independent of other works, a brief account of the special theory of relativity and of the general theory of Hamiltonian dynamics is included in the first two chapters.

Only the periodic phenomena are considered in this first volume, and a few easy applications capable of illustrating the general principles are also discussed. It is intended to present the transformation theory, quantum statistics, theory of aperiodic phenomena, quaternionic mechanics, field theory and higher applications in a second volume.

For further study of the subject, the reader is referred to the original papers and treatises, a list of which is given in the bibliography.

I am very much grateful to my respected teacher Professor W. Heisenberg, who went through the manuscript with me, and made many valuable suggestions. I am also indebted to my friend Professor N. R. Sen for much helpful discussion.

RAZIUDDIN SIDDIQI.

Hyderabad-Deccan,
April 1937.

TABLE OF CONTENTS

CHAPTER I

THE THEORY OF RELATIVITY

SECTION PAGE

1·1 The Michelson-Morley Experiment 1
1·2 Einstein's Postulates 3
1·3 The Lorentz-Transformation 6
1·4 Some Consequences of the Lorentz-Transformation .. 11
1·5 The Four-Dimensional Continuum 13
1·6 Relativistic Mechanics 15
1·6 (1) Mass and Momentum 16
1·6 (2) Relativistic Equations of Motion and Energy .. 19
1·7 Minkowski's Four-Dimensional Vectors 21

CHAPTER II

GENERAL THEORY OF DYNAMICS

2·1 Hamilton's Variation Principle 25
2·2 Lagrange's Equations 29
2·3 Hamilton's Canonical Equations 31
2·4 Relativistic Lagrangian and Hamiltonian 33
2·5 Canonical Transformations. Hamilton-Jacobi Equation 35
2·6 Integration of the Canonical Equations 38
2·7 Periodic Motions and the Action and Angle Variables 41

CHAPTER III

STRUCTURE OF THE ATOM

3·1 Rutherford's Investigations 45
3·2 Rutherford's Atom-Model 48
3·3 The Neutron 50
3·4 The Positron 51
3·5 Structure of the Nucleus 51

CHAPTER IV

THE QUANTUM THEORY OF RADIATION

SECTION PAGE

4·1 The Dual Nature of Light. Einstein's Light-Quantum 53
4·2 Relativity Theory of the Photon. Derivation of Planck's Law 54
4·3 Black-body Radiation, and Planck's Quantum Hypothesis 58
4·4 Theory of the Photo-electric Effect 61
4·5 Theory of the Compton Effect 63

CHAPTER V

BOHR'S QUANTUM THEORY OF THE ATOM

5·1 Classical Theory of the Hydrogen Atom .. 69
5·2 Bohr's Postulates 71
5·3 Bohr's Theory of the Hydrogen Atom 72
5·4 Motion of the Nucleus 75
5·5 Generalisation of the Quantum Condition 78
5·6 The Correspondence Principle 80
5·7 Relativity Theory of the Fine Structure 83
5·8 Development of Bohr's Theory 94
5·8 (1) Atomic Magnetism. Magnetic Moment of the Atom 95
5·8 (2) The Spectra of Heavier Atoms 97
5·8 (3) Pauli's Exclusion Principle 100
5·9 Bohr's Theory of the Structure of Atoms, and the Periodic Table of Chemical Elements 101

CHAPTER VI

HEISENBERG'S QUANTUM MECHANICS

6·1 Critique of Bohr's Theory 105
6·2 Heisenberg's Theory 106
6·3 The Algebra of Matrices 110
6·4 Poisson Brackets 116

SECTION PAGE
6·5 The Quantum Conditions and the Equations of Motion 121
6·6 Equation of Energy. Bohr's Frequency Relation .. 124
6·7 The Harmonic Oscillator 126
6·8 (1) The Angular Momentum 133
6·8 (2) Selection and Polarisation Rules 136
6·9 (1) Further Development of Matrix Algebra 137
6·9 (2) Matrix Analysis 140
6·9 (3) Transformation of a Matrix to a Diagonal Matrix. Determination of Eigen-values 144
6·9 (4) Canonical Transformations 146

CHAPTER VII

THE WAVE THEORY OF MATTER

7·1 De Broglie's Conception of Matter-Waves 148
7·2 De Broglie's Theory 149
7·3 Wave-Group and Group Velocity 153
7·4 The Analogy between Mechanics and Optics .. 154
7·5 Schrödinger's Wave Equation 156
7·6 Experiments on the Diffraction of Electrons .. 157

CHAPTER VIII

WAVE MECHANICS

8·1 Eigen-values and Eigen-functions 161
8·2 Polynomial Method for Determining the Eigen-values 162
8·3 Bechert's Method for Determining the Eigen-values .. 166
8·4 The Linear Harmonic Oscillator 168
8·4 (1) The Linear Harmonic Oscillator (Polynomial Method) 170
8·5 The Hydrogen-like Atom 172
8·6 Equivalence of Matrix Mechanics and Wave-Mechanics 178
8·6 (1) The Wave-Equation 179
8·6 (2) Derivation of Matrix Elements from Eigen-functions 181

CHAPTER IX

PERTURBATION THEORY

SECTION PAGE

9·1 Outline of the Perturbation Method 188
9·2 The Vibrating Membrane 189
9·3 Perturbation Theory for the Non-Degenerate Case .. 191
9·4 The Atom under an Electric Field. The Stark-effect 197
9·5 Perturbation Theory for the Degenerate Case .. 204
9·6 Theory of the Neutral Helium Atom 209
9·7 Systems Containing Similar Particles 215
9·8 Perturbation Theory in Matrix Mechanics.. .. 217

CHAPTER X

PHYSICAL INTERPRETATIONS AND HEISENBERG'S PRINCIPLE OF INDETERMINACY

10·1 (1) Physical Significance of the Wave-Function Schrödinger's Interpretation 225
10·1 (2) Born's Interpretation 227
10·2 Calculation of Intensities. Selection Rules .. 229
10·3 Complimentary Character of the Corpuscular and the Wave Conceptions 232
10·4 Determination of the Position of a Free Electron .. 234
10·5 Determination of the Velocity of an Electron .. 237
10·6 The Uncertainty Relation. Theory of Operators .. 239
10·7 Principle of Indeterminacy and the Concept of Causality 245

CHAPTER XI

RELATIVISTIC QUANTUM MECHANICS

11·1 Relativistic Electrodynamics 248
11·1 (1) Conservation of Electric Charge 248
11·1 (2) Density of Electric Charge and Electric Current .. 248
11·1 (3) The Maxwell-Lorentz Field Equations 250
11·1 (4) Relativistic Dynamics of a Charged Particle in an Electromagnetic Field 254

SECTION PAGE

11·2 Earlier Attempts at a Relativistic Quantum Theory 258

11·3 Dirac's Relativistic Theory for a Free Electron .. 261

11·4 The Electron under an Arbitrary Field 263

11·5 Existence of the Magnetic Moment of the Electron .. 264

11·6 Existence of the Spin Angular Momentum of the Electron 267

11·7 Theory of the Hydrogen Atom. Fine Structure of Spectral Lines 270

11·8 The Atom under an External Magnetic Field. The Normal Zeeman Effect 278

BIBLIOGRAPHY 283

INDEX 289

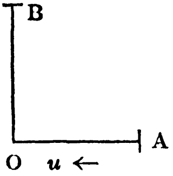
CHAPTER I.

THE THEORY OF RELATIVITY.

1·1. *The Michelson-Morley Experiment.*

THE starting point of the relativity theory was the experiment of Michelson and Morley in 1887 to measure the speed of the earth's motion through the ether by means of sending light signals simultaneously in two directions at right angles to each other, and measuring the times they take to come back to their starting point, after they are reflected at equal distances.

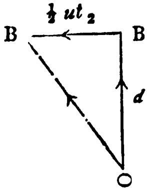
Suppose the velocity of the light signals is c, and that of the earth through the ether is u. Suppose also that the earth is moving from east to west.



Let O be the starting point of the signals, and let OA be any distance d towards east, and OB an equal distance d towards north ; and let there be two mirrors at A and B as shown, at which the light signals are reflected. The earth is moving in the direction shown by the arrow. The signal along OA will have the velocity $c+u$, and therefore the time it will take to reach A will be $\frac{d}{c+u}$. While coming back, its velocity will be $c-u$, and the time spent in arriving at O will be $\frac{d}{c-u}$. Thus the whole time from O to A and back to O is

$$\begin{aligned} t_1 &= \frac{d}{c+u} + \frac{d}{c-u} = \frac{2\,cd}{c^2-u^2} \\ &= 2\frac{d}{c}\frac{1}{1-\frac{u^2}{c^2}}. \end{aligned} \tag{1}$$

Consider next the signal to OB and back, and suppose it takes time t_2 to do the whole journey. During this time when the signal arrives back to O, the whole apparatus would have moved a distance ut_2 westwards. Therefore, during the outward journey alone, the distance travelled by the apparatus is half, *i.e.*, $\frac{1}{2}ut_2$.



Thus the actual path of the signal is the hypotenuse OB. So

$$OB = \sqrt{d^2 + \tfrac{1}{4}u^2t_2^2}.$$

This distance is traversed with the velocity c in time $\frac{1}{2}t_2$, so that

$$\tfrac{1}{2}ct_2 = \sqrt{d^2 + \tfrac{1}{4}u^2t_2^2},$$

which gives, on squaring and solving for t_2^2:

$$t_2^2 = \frac{4d^2}{c^2 - u^2} = \frac{4d^2}{c^2}\,\frac{1}{1 - \dfrac{u^2}{c^2}},$$

i.e.,

$$t_2 = 2\frac{d}{c}\,\frac{1}{\sqrt{1 - \dfrac{u^2}{c^2}}}\,. \tag{2}$$

Thus it is evident that $t_1 \neq t_2$ unless $\frac{u}{c} = 0$. The difference $t_1 - t_2$, if observed, would give us the value u of the earth's velocity through the ether. But when Michelson and Morley performed their experiment, they got the completely unexpected result that both the signals arrive at O at the same time, *i.e.*, $t_1 = t_2$. The experiment was performed at all times of the year, and with very delicate instruments, but the result was consistently the same, *viz.*, $t_1 = t_2$. And this in spite of the fact that neither u is zero, nor c is ∞.

To explain this surprising result, Fitz Gerald and H. A. Lorentz, independently in 1892, suggested that the distance OA in the direction of the earth's motion contracts automatically, whereas the distance OB in the direction at right angles to the direction of the earth's motion remains unaltered. Thus the signal along OA and back, which has to travel a shorter distance, makes the journey to and fro in the same time as the signal along OB and back.

If we accept this hypothesis, we can calculate the contraction factor very easily.

Suppose the length OA is shortened to kd, where k is the contraction factor. Then the journey OA takes the time $\frac{kd}{c+u}$, and the journey AO takes the time $\frac{kd}{c-u}$. Thus the whole journey to and fro occupies the time,

$$t_1{}' = \frac{kd}{c+u} + \frac{kd}{c-u} = \frac{2kd}{c}\left(\frac{1}{1-\frac{u^2}{c^2}}\right). \tag{3}$$

If $t_1{}' = t_2$, then equating the right-hand sides of (2) and (3) we get

$$\frac{k}{1-\frac{u^2}{c^2}} = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}},$$

or

$$k = \sqrt{1-\frac{u^2}{c^2}}. \tag{4}$$

Thus, according to Fitz Gerald and Lorentz, if a body is travelling through ether with a velocity u, each unit of length in the direction of its motion is contracted to $k = \sqrt{1-\frac{u^2}{c^2}}$, where c is the velocity of light. This is called the Fitz Gerald-Lorentz contraction, and it explains the Michelson-Morley experiment. But the contraction itself remained a puzzle, until Einstein published his theory of relativity in 1905.

1·2. *Einstein's Postulates.*

From the failure of the Michelson-Morley experiment, and all other experiments of the same kind, Einstein argued that no experiment can ever detect the earth's speed through the ether. He went a step further, and proclaimed that Nature itself does not know of any motion of the earth through the ether. There is no such thing as absolute motion; all movements in Nature are relative. Consequently, all axes of co-ordinates, or "frames of reference", as they are now called, moving with the various observers, have the same status. The physical laws should, therefore, be independent of the motion of the observer.* At first Einstein confined himself to uniform motions of translation only, and put forward the special, or restricted, theory of relativity. Later, in 1915, he generalised his theory to embrace any kind of accelerated motion.

The problem, therefore, which Einstein set himself was. "How must the Laws of Nature be framed in order that it may be impossible to make an essential distinction between two systems which are in relative motion."

This idea is found to a certain extent in Newtonian mechanics also. Consider, *e.g.*, a system of particles $m_1, m_2, .., m_n$, moving under their mutual attractions.

Let (x_r, y_r, z_r) be the coordinates of the r^{th} particle in a co-ordinate system K, and let r_{12} denote the distance between the first and the second particle. The potential ϕ of the system will then be a function of $r_{12}, r_{13}, .., r_{23}, \ldots..$

The equations of motion in the co-ordinate system K will be

$$m_r \frac{d^2x_r}{dt^2} = -\frac{\partial\phi}{\partial x_r}, \text{ etc.} \tag{1}$$

Now we take another co-ordinate system K′ (x', y', z') of which the origin O moves with a constant velocity v relative to the system K in the direction of the x-axis. If initially both origins are supposed coincident, the distance between them after the lapse of time t will be vt, so that the new co-ordinates of the particles will become,

$$x_r' = x_r - vt, \qquad y_r' = y_r, \qquad z_r' = z_r. \tag{2}$$

However, the mutual distance of the particles will remain unaltered, so that,

$$r_{12}' = r_{12}, \text{ etc.} \tag{3}$$

Since according to the classical ideas, time is absolute, we have $t' = t$, and the equations of motion in K′ are given by

$$m_r \frac{d^2x_r'}{dt^2} = -\frac{\partial\phi'}{\partial x_r'}, \text{ etc.} \tag{4}$$

where

$$\phi' = \phi\,(r_{12}', r_{13}', \cdots, r_{23}', \cdots), \tag{5}$$

showing that the new equations of motion (4) are of the same form as the old equations of motion (1).

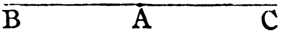
The transformation (2) is called the " Galilean-transformation ", and we have seen that from the point of view of mechanics the two co-ordinate systems K and K′ connected by the relation (2) are equivalent. But they are not equivalent for every phenomenon of nature. We shall give an example from electromagnetic phenomena to show that two co-ordinate systems connected by a Galilean-transformation are not equivalent.

Suppose a material point A lies at rest at the origin O in a co-ordinate system K, and sends at time $t = 0$ a short light signal in all directions.

At a subsequent time t the points which receive the signal lie on a sphere whose equation is given by:

$$x^2 + y^2 + z^2 - c^2t^2 = 0. \tag{6}$$

Suppose there are two more material points B and C which lie on a line through A parallel to the x-axis, and at equal distances from it on either side.

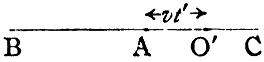


The material points B and C are also at rest in K, and will receive the signal simultaneously, *i.e.*, for the same value of t.

Now relative to the system K′, the material points A, B, C have the velocity v parallel to CB. The two origins O and O′ coincide initially, so that in the system K′ the light signal is sent at time $t' = 0$ from the material point A at O′. After time t', all the points which receive the signal will lie on the sphere,

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \tag{7}$$

Meanwhile A would have travelled a distance vt' from the origin O′, so that the signal will arrive at C earlier than at B.



There is no real value of t' for which the two points B and C will lie on the sphere (7). Thus we see that two events which are simultaneous in the system K are not simultaneous in the system K′.

Einstein was thus led to the conclusion that the classical notion of an absolute universal time cannot be retained. Each observer in relative motion has his own special or " local " time, as it is called, which depends on his motion. We must find, therefore, a transformation connecting the co-ordinates of the two systems K and K′, which is capable of replacing the Galilean-transformation. Such a transformation was first discovered by Lorentz in 1904 on electromagnetic grounds. But its true significance was recognised later (1905) by Einstein who derived it on broad principles which revolutionized the then prevalent philosophy of space, time and motion.

Based on the experiments of Michelson and Morley, Trouton and Noble, and others, Einstein put forward the following two fundamental postulates, called the " Postulate of Relativity " and the " Postulate of the Constancy of the Velocity of Light ".

The Postulate of Relativity.—It is impossible by any physical measurement to trace an essential distinction between any two co-ordinate systems which are in uniform relative motion to each other.

The Postulate of the Constancy of the Velocity of Light.—The velocity of light is constant in all directions, and in all co-ordinate systems which are in uniform relative motion to each other.

In the next section we shall derive the Lorentz-transformation, using these two postulates.

1 · 3. *The Lorentz-Transformation.*

Let the two frames of reference be K with the co-ordinates of space and time x, y, z, t, and K′ with the co-ordinates x', y', z', t'. We assume the two origins to be coincident initially, *i.e.*, to the world-point $x = y = z = t = 0$ in K corresponds the world-point $x' = y' = z' = t' = 0$ in K′.

We assume further that the corresponding co-ordinate axes in both frames are parallel and that K′ moves with uniform velocity v relative to K in the direction of the x-axis. This restriction to uniform translation is characteristic of the special theory of relativity, with which alone we shall be concerned in this book. The consideration of any kind of accelerated motion is the standpoint of the "General" Theory of Relativity.

Thus we assume that each point of K′ moves with the uniform velocity v relative to K, and that, inversely, each point of K moves with the uniform velocity $(-v)$ relative to K′; that v is parallel to the x-axis and that the x-axis and the x'-axis are always coincident.

Now it must be taken as an obvious law of Nature that no one point in any frame should have privileged position, but that all points of space and moments of time should have the same status. This means that finite values of co-ordinates in K must remain finite in K′ also. Further, it is evident that to uniform and straight motions in K must correspond uniform and straight motions in K′. All this makes it necessary that our equations of transformation must be linear. For, if the equations were non-linear, the origin will have a physically privileged position in at least one of the two frames. Besides, the uniform and straight character of the motion will not be preserved.

The coincidence of the x-axis and the x'-axis is only then possible when from $y=0$, $z=0$, we get $y'=0$, $z'=0$. Therefore, the transformation for y and z must be of the form

$$y'=\alpha y+\beta z, \qquad z'=\gamma y+\delta z,$$

where α, β, γ, δ may depend on the velocity v. But pure rotations in space are inessential, and therefore we assume that the x-y plane ($z=0$) transforms into the x'-y' plane ($z'=0$). Thus y' can depend only on y and z' can depend only on z, so that

$$y'=\epsilon(v)\,y, \qquad z'=\epsilon'(v)\,z.$$

But motion takes place in the x direction only, so that the y- and z-directions being both at right angles to it have equal status. ϵ and ϵ' must therefore be equal, and the transformations for y and z become

$$y'=\epsilon(v)\,y, \qquad z'=\epsilon(v)\,z. \tag{1}$$

We can show further that ϵ is equal to unity. In fact, it is clear that $\epsilon(v)=\epsilon(-v)$, because it is immaterial for the y- and z-directions whether K′ moves relative to K in the direction of the positive x-axis or the negative x-axis.

Now since inversely K moves with velocity $(-v)$ relative to K′, the inverse transformation should be

$$y=\epsilon(-v)\,y', \qquad z=\epsilon(-v)\,z',$$

or, on account of $\epsilon(v)=\epsilon(-v)$,

$$y=\epsilon(v)\,y', \qquad z=\epsilon(v)\,z'. \tag{2}$$

From (1), however, we get for the inverse transformation

$$y=\frac{1}{\epsilon(v)}\,y', \quad z=\frac{1}{\epsilon(v)}\,z'. \tag{3}$$

Comparing (2) and (3) we get

$$\frac{1}{\epsilon(v)}=\epsilon(v), \quad \text{or } \epsilon(v)=1. \tag{4}$$

Thus the transformations for y and z are finally

$$y'=y, \qquad z'=z. \tag{5}$$

It remains for us to determine the transformation for x and t. By hypothesis, the point $x'=0$ must move with velocity v along the positive x-axis, *i.e.*, $x'=0$ must be identical with $x=vt$. Similarly $x=0$ must be equivalent to $x'=-vt'$. The required transformation must therefore have the form:

$$\begin{aligned} x' &= \gamma'(x-vt), \\ x &= \gamma(x'+vt'). \end{aligned} \tag{6}$$

It is easily seen that γ' must be equal to γ, if the relativity postulate is to hold. For, suppose an observer S′ in K′ measures the length of a rod which is at rest in K, and which has a length l as measured by an observer S in K. Then suppose the rod moves with K′, and its length is measured by S. According to the relativity postulate the two measurements must give the same length.

In the first measurement the two ends of the rod are given by $x = 0$ and $x = l$ in K. In K′ these two points have at time $t' = 0$ the co-ordinates $x' = 0$ and $x' = \frac{1}{\gamma} l$. For S′ the rod has therefore the length $\frac{l}{\gamma}$, so that a unit rod will have the length $\frac{1}{\gamma}$.

Now in the second measurement the rod is at rest in K′, and its two ends are given by $x' = 0$ and $x' = l$. As measured by S, these two points have at $t = 0$ the co-ordinates $x = 0$ and $x = \frac{l}{\gamma'}$. For S the rod has therefore the length $\frac{l}{\gamma'}$, so that a unit rod will have a length $\frac{1}{\gamma'}$.

By the relativity postulate, therefore, we must have

$$\gamma' = \gamma. \tag{7}$$

To determine the value of γ we use the postulate of the constancy of the velocity of light. Suppose a light signal is given at origin O at the time $t = 0$, $t' = 0$, and that it is received at a point situated anywhere on the x-axis. This event would be described by S by a relation between x and t, and by S′ by a relation between x' and t'. In any case we must have, of course, $x = ct$, and $x' = ct'$, because, according to the postulate, c is the velocity of light in both systems. Substituting these values in the equations (6), we get

$$ct' = \gamma t\ (c - v),$$
$$ct = \gamma t'\ (c + v).$$

Multiplying both sides of these two equations, we find

$$c^2 t' t = \gamma^2 t' t\ (c^2 - v^2),$$

or,

$$\gamma^2 = \frac{c^2}{c^2 - v^2} = \frac{1}{\dfrac{c^2 - v^2}{c^2}} = \frac{1}{1 - \dfrac{v^2}{c^2}}. \tag{8}$$

Thus

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{9}$$

where we have taken the positive sign of the square root because for $v = 0$, K′ must be identical with K.

From the two equations (6), we can find t' in terms of t and x. Thus

$$vt' = \frac{x}{\gamma} - x' = \frac{x}{\gamma} - \gamma(x - vt) = \frac{x}{\gamma} - \gamma x + \gamma vt,$$

so that

$$\begin{aligned} t' &= \frac{1}{v}\left\{\frac{x}{\gamma}(1 - \gamma^2) + \gamma vt\right\} = \gamma\left\{t + \frac{x}{v}\left(\frac{1}{\gamma^2} - 1\right)\right\} \\ &= \gamma\left\{t + \frac{x}{v}\left(-\frac{v^2}{c^2}\right)\right\} = \gamma\left(t - \frac{xv}{c^2}\right). \end{aligned} \tag{10}$$

From (6) and (10) we have on account of (9) and (5)

$$\begin{aligned} x' &= \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad y' = y, \quad z' = z, \\ t' &= \frac{t - x\frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}. \end{aligned} \tag{11}$$

These are the required equations of the Lorentz-transformation. For the inverse transformation we get easily :

$$\begin{aligned} x &= \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad y = y', \quad z = z', \\ t &= \frac{t' + \frac{v}{c^2}x'}{\sqrt{1 - \frac{v^2}{c^2}}}. \end{aligned} \tag{12}$$

For the limiting value $c \to \infty$, the Lorentz-transformation (11) and the Galilean-transformation (2) § 1·2 are identical, as we should expect.

We can see now that the two frames of reference K and K′ connected by the Lorentz-transformation are equivalent for electromagnetic phenomena. We consider the law of propagation of light.

A spherical wave starting from O at time $t = 0$ has reached the spherical surface $x^2 + y^2 + z^2 - c^2t^2 = 0$, as observed by S. The same event must be expressible in K′ by the equation $x'^2 + y'^2 + z'^2 - c^2t'^2 = 0$. Now, we have from (11):

$$x'^2 - c^2t'^2 = \gamma^2 (x - vt)^2 - c^2\gamma^2 \left(t - \frac{v}{c^2}\, x\right)^2$$

$$= x^2 \left(\gamma^2 - \gamma^2 \frac{v^2}{c^2}\right) + t^2 (\gamma^2 v^2 - c^2\gamma^2) + 2xt\,(\gamma^2 v - \gamma^2 v).$$

But

$$\gamma^2 - \gamma^2 \frac{v^2}{c^2} = \gamma^2 \left(1 - \frac{v^2}{c^2}\right) = \frac{1}{1 - \frac{v^2}{c^2}} \left(1 - \frac{v^2}{c^2}\right) = 1,$$

$$\gamma^2 v^2 - c^2\gamma^2 = - c^2 \left(\gamma^2 - \gamma^2 \frac{v^2}{c^2}\right) = - c^2,$$

so that

$$x'^2 - c^2t'^2 = x^2 - c^2t^2.$$

Thus we have finally

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2. \tag{13}$$

We get the same result also in the differential form. The equation of wave-motion in the frame K is

$$\Box\phi \equiv \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\phi = 0. \tag{14}$$

This equation is just the mathematical formulation of the law of propagation of light in vacuum. If we wish to transform (14) to the frame K′, we must substitute in the function $\phi\,(x, y, z, t)$ the values from (12). Then we have

$$\frac{\partial\phi}{\partial x'} = \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial x'} + \frac{\partial\phi}{\partial t}\frac{\partial t}{\partial x'} = \gamma\frac{\partial\phi}{\partial x} + \gamma\frac{v}{c^2}\frac{\partial\phi}{\partial t};$$

$$\frac{\partial^2\phi}{\partial x'^2} = \left(\gamma\frac{\partial}{\partial x} + \gamma\frac{v}{c^2}\frac{\partial}{\partial t}\right)\left(\gamma\frac{\partial\phi}{\partial x} + \gamma\frac{v}{c^2}\frac{\partial\phi}{\partial t}\right)$$

$$= \gamma^2\frac{\partial^2\phi}{\partial x^2} + 2\gamma^2\frac{v}{c^2}\frac{\partial^2\phi}{\partial x\partial t} + \gamma^2\frac{v^2}{c^4}\frac{\partial^2\phi}{\partial t^2};$$

$$\frac{\partial^2\phi}{\partial y'^2} = \frac{\partial^2\phi}{\partial y^2};\quad \frac{\partial^2\phi}{\partial z'^2} = \frac{\partial^2\phi}{\partial z^2};$$

$$\frac{\partial\phi}{\partial t'} = \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial t'} + \frac{\partial\phi}{\partial t}\frac{\partial t}{\partial t'} = \gamma v\frac{\partial\phi}{\partial x} + \gamma\frac{\partial\phi}{\partial t};$$

$$\frac{\partial^2\phi}{\partial t'^2} = \left(\gamma v\frac{\partial}{\partial x} + \gamma\frac{\partial}{\partial t}\right)^2\phi,$$

$$= \gamma^2 v^2\frac{\partial^2\phi}{\partial x^2} + 2\gamma^2 v\frac{\partial^2\phi}{\partial x\partial t} + \gamma^2\frac{\partial^2\phi}{\partial t^2}.$$

Therefore

$$\begin{aligned}\Box'\phi &= \left(\frac{\partial^2}{\partial x'^2}+\frac{\partial^2}{\partial y'^2}+\frac{\partial^2}{\partial z'^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\right)\phi\\ &= \gamma^2\frac{\partial^2\phi}{\partial x^2}+2\gamma^2\frac{v}{c^2}\frac{\partial^2\phi}{\partial x\partial t}+\gamma^2\frac{v^2}{c^4}\frac{\partial^2\phi}{\partial t^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}\\ &\qquad -\gamma^2\frac{v^2}{c^2}\frac{\partial^2\phi}{\partial x^2}-2\gamma^2\frac{v}{c^2}\frac{\partial^2\phi}{\partial x\partial t}-\gamma^2\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}.\end{aligned}$$

Thus

$$\begin{aligned}\Box'\phi &= \left(\gamma^2-\gamma^2\frac{v^2}{c^2}\right)\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}-\frac{1}{c^2}\left(\gamma^2-\frac{\gamma^2v^2}{c^2}\right)\frac{\partial^2\phi}{\partial t^2}\\ &= \Box\phi.\end{aligned}$$

The characteristic expression $\Box\phi$ for the propagation of light is therefore invariant for the Lorentz-transformation.

1·4. *Some Consequences of the Lorentz-Transformation.*

(*a*) *The rule for composition of velocities.*—Consider a particle moving along the x-axis, whose velocity measured by S is u, so that $u=\frac{dx}{dt}$. Then its velocity measured by S′ would be

$$\begin{aligned}u' &= \frac{dx'}{dt'}=\frac{\gamma(dx-vdt)}{\gamma\left(dt-\frac{v}{c^2}dx\right)}=\frac{\frac{dx}{dt}-v}{1-\frac{v}{c^2}\frac{dx}{dt}}\\ &= \frac{u-v}{1-\frac{uv}{c^2}}. \qquad (1)\end{aligned}$$

In Newtonian mechanics, if S′ is moving with a velocity v, relative to S, and if the particle is moving with a velocity u relative to S, then the velocity of the particle relative to S′ will be simply $u-v$. The theory of relativity increases this velocity to the amount given by the expression (1).

The expression (1) for the composition of two velocities is true for all $u<c$. But if $u=c$, then

$$u'=\frac{c-v}{1-\frac{cv}{c^2}}=c. \qquad (2)$$

Thus when a particle is moving with the velocity c of light, as observed by S, its velocity as observed by S′ is also c. The motion of the frame of reference of S′ has no effect on the velocity of light. This verifies that the Lorentz-transformation equations are consistent with the fundamental postulate of the theory, and consequently with the results of the Michelson-Morley and other experiments.

(*b*) *Contraction in length.*—Let a rigid rod of length l_0 be fixed in the frame K′. We wish to find its length as measured by S.

Suppose first that the rod lies parallel to the direction Oy'. If the two ends of the rod have co-ordinates y_1', y_2' respectively, then

$$l_0 = y_2' - y_1'.$$

For the observer S the two ends will have co-ordinates y_1, y_2 respectively, and therefore the length l of the rod as measured by S will be

$$l = y_2 - y_1 = y_2' - y_1' = l_0. \tag{3}$$

We see, therefore, that lengths remain unaltered in a direction perpendicular to the direction of motion of the body as measured by an observer S.

Suppose now that the rod lies parallel to the direction Ox'. As shown in § 1·3, its length as measured by S would be $\frac{l_0}{\gamma'}$. But from (7) and (9) § 1·3, we know that

$$\gamma' = \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{4}$$

Thus the length l of the rod as measured by S would be

$$l = \frac{l_0}{\gamma} = \sqrt{1 - \frac{v^2}{c^2}}\, l_0 = kl_0, \tag{5}$$

where k is the contraction factor introduced in (4) § 2·1. This proves that each unit of length of the rod in the direction of its motion is contracted to k when measured by an observer S. Thus the Fitz Gerald-Lorentz contraction is a natural consequence of the relativity theory.

(*c*) *Dilatation in time.*—Again let t_1' and t_2' be two times registered by a clock fixed in the frame of reference K′ at the point x', and let t_1 and t_2 be the corresponding times registered by a clock fixed in K. Then from (12) § 1·3, we have

$$t_1 = \gamma\left(t_1' + \frac{vx'}{c}\right), \quad t_2 = \gamma\left(t_2' + \frac{vx'}{c^2}\right). \tag{6}$$

Now the interval of time τ_0 as measured by S′ is·

$$\tau_0 = t_2' - t_1', \tag{7}$$

and the corresponding interval τ as measured by S is

$$\tau = t_2 - t_1.$$

But from (6) we get on subtraction

$$t_2 - t_1 = \gamma\,(t_2' - t_1'), \tag{8}$$

so that from (7) and (8) we have

$$\tau = \gamma\tau_0 = \frac{\tau_0}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{9}$$

Hence each unit interval of time as registered by a clock fixed in K′ appears to the observer S to be lengthened to an interval γ. In other words, a clock moving relatively to the frame K, appears to an observer S at rest in K to be going slow.

It can be proved similarly that rods fixed in K appear to S′ to be contracted by the same amount, and that clocks fixed in K appear to S′ to be going slow. There is thus a complete reciprocity between the observers S and S′ as regards the contraction in length and dilatation in time.

1 · 5. *The Four-Dimensional Continuum, and the Interval between a pair of Events.*

Elaborating Einstein's ideas, H. Minkowski pointed out in 1908 that our external world is composed " not of *points* in the Euclidean 3-dimensional space but of *events* in a 4-dimensional space-time continuum ". An event can therefore be considered to have four co-ordinates (x_1, x_2, x_3, x_4) of which three, say x_1, x_2, x_3, are space co-ordinates, and the fourth, x_4, involves time.

Let (x_1, x_2, x_3, x_4) and $(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3, x_4 + dx_4)$ be the co-ordinates of two neighbouring events in any frame of reference. Generalising the idea of distance between two neighbouring points in Euclidean space, we assume that in the special theory of relativity the " interval " between two neighbouring events in the 4-dimensional continuum is given by the normal quadratic form

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2. \tag{1}$$

Now consider all the events for which x_4 has some constant specified value; these will form a 3-dimensional continuum. Since dx_4 is zero for every pair of these events, their mutual interval is given by writing $dx_4 = 0$ in (1). Thus

$$ds^2 = dx_1^2 + dx_2^2 + dx^2_3. \tag{2}$$

This is the same as the Cartesian formula for the distance between two neighbouring points in Euclidean space, with x_1, x_2, x_3 written in place of the usual x, y, z. Thus in general we shall have

$$ds^2 = dx^2 + dy^2 + dz^2 + dx_4^2, \tag{3}$$

where the co-ordinate x_4 involves the time.

Now in Euclidean geometry, we know that the distance between two points is invariant for any change of axes. Similarly, in our 4-dimensional geometry we must have the interval between two events independent of the transformation from one frame of reference to another. Our equations of transformations are, however, not the Galilean or Cartesian equations, but the Lorentz-equations (11) § 1·3. As proved in (13) § 1·3, it is evident that the quadratic differential expression which is invariant for the Lorentz-transformations is $dx^2 + dy^2 + dz^2 - c^2dt^2$. Thus the invariant interval between a pair of neighbouring events must be given by

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2. \tag{4}$$

Comparing (3) and (4), we see that the fourth co-ordinate x_4 involves the time according to the formula

$$x_4 = ict, \ i = \sqrt{-1}. \tag{5}$$

Thus the four co-ordinates of an event are

$$x_1 = x, \ x_2 = y, \ x_3 = z, \ x_4 = ict, \tag{6}$$

where c denotes the velocity of light in vacuum.

The formula (4) for the interval is the basis of the whole special theory of relativity.

The "eigentime".—The "eigentime" or "local time" of a body is defined to be the time τ shown by the clock which is moving with the body. The ordinary time t is measured from the system at rest.

Suppose the body fixed in the frame K′, and consequently moving with velocity v relative to the frame K. Then the eigentime τ is the same which we have called t' hitherto. Let two neighbouring events have times τ, $\tau + d\tau$ according to the clock moving with the body. The frame K gives times t, $t + dt$ for these two events. Now, even if the velocity v is not uniform, we can suppose that for small intervals of time it can be approximated by a rectilinear and constant velocity. Then we can use the results proved for dilatation of time in (*c*) § 1·4, if we write $d\tau$ for τ_0 and dt for τ. Thus

$$dt = \gamma d\tau = \frac{d\tau}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{7}$$

The observer S finds the time-interval between the two events greater than the eigentime-interval.

From (7) we get on squaring,

$$d\tau^2 = dt^2 \left(1 - \frac{v^2}{c^2}\right) = \frac{1}{c^2} (c^2dt^2 - v^2dt^2). \tag{8}$$

But the instantaneous velocity v of the body, as measured by S, is given by

$$v^2 = \frac{dx^2 + dy^2 + dz^2}{dt^2},$$

or

$$v^2\, dt^2 = dx^2 + dy^2 + dz^2. \tag{9}$$

Substituting (9) in (8) we get :

$$d\tau^2 = \frac{1}{c^2} (c^2dt^2 - dx^2 - dy^2 - dz^2). \tag{10}$$

The right-hand side of (10) is invariant for the Lorentz-transformations. The eigentime, therefore, is also invariant.

1·6. *Relativistic Mechanics.*

According to Einstein's principle of relativity, all laws of Nature must be invariant for the Lorentz-transformations. The Newtonian equations of motion do not possess this character and must therefore be replaced by others which satisfy the principle. While doing this we must remember that for velocities very small compared to the fundamental velocity c, Einstein's equations must tend to Newton's equation.

In Newtonian mechanics, suppose that a particle of mass m, moving in a given field of force, is situated at time t at a point

(x, y, z). The components of its velocity vector $\vec{u}$ are $(\dot{x}, \dot{y}, \dot{z})$,[1] and those of its momentum vector $\vec{p}$ are

$$p_x = m\dot{x},\ p_y = m\dot{y},\ p_z = m\dot{z}. \tag{1}$$

If $\vec{F}$ is the Newtonian force-vector, the equations of motion are

$$\frac{dp_x}{dt} = F_x,\ \frac{dp_y}{dt} = F_y,\ \frac{dp_z}{dt} = F_z. \tag{2}$$

Multiplying these equations by $u_x = \dot{x}$, $u_y = \dot{y}$, $u_z = \dot{z}$ respectively, and adding, we get the energy-equation

$$\frac{d}{dt}\left(\tfrac{1}{2}mu^2\right) = (\vec{F}\vec{u}) = (F_x u_x + F_y u_y + F_z u_z), \tag{3}$$

showing that the rate of change of kinetic energy is equal to the work done by the external forces in unit time.

1·6 (1). *Mass and Momentum.*

Now in Newtonian mechanics, the mass of a moving body is always the same constant however the body may be moving, and with reference to whatever frame the mass is measured. From the equation (2) § 1 · 6 it is evident further that if $\vec{F} = 0$, *i.e.*, if no external forces are acting on the body, then the momentum remains also unaltered.

In building up the relativistic mechanics, we naturally wish to take over these laws. But it is found that both these laws are not compatible in the new theory. If we assume the law of conservation of momentum to hold, then we are forced to the conclusion that the mass of a body is not invariant, but changes with its velocity in a prescribed manner.

Let m_1, m_1' be the mass of a body and u_1, u_1' its velocity in the x-direction as measured by the two observers S and S′ respectively, where it is assumed, as before, that the frame of reference K′ is moving with respect to the frame K with the velocity v in the x-direction.

We write

$$\gamma_1 = \frac{1}{\sqrt{1 - \frac{u_1^2}{c^2}}},\quad \gamma_1' = \frac{1}{\sqrt{1 - \frac{u_1'^2}{c^2}}},\quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{1}$$

---

[1] We shall often denote by dots the total derivatives with respect to time.

Then from equation (1), § 1·4 we have

$$u_1' = \frac{u_1 - v}{1 - \frac{u_1 v}{c^2}}. \tag{2}$$

Therefore

$$\begin{aligned}\gamma\,\gamma_1\,(u_1 - v) &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \cdot \frac{1}{\sqrt{1 - \frac{u_1^2}{c^2}}}\,(u_1 - v) \\ &= \frac{u_1 - v}{\sqrt{1 - \frac{1}{c^2}(v^2 + u_1^2) + \frac{u_1^2 v^2}{c^4}}}.\end{aligned} \tag{3}$$

Also from (1) and (2) we get

$$\gamma_1' u_1' = \frac{u_1 - v}{\sqrt{\left(1 - \frac{u_1'^2}{c^2}\right)\left(1 - \frac{u_1 v}{c^2}\right)^2}}. \tag{4}$$

But

$$1 - \frac{u_1'^2}{c^2} = 1 - \frac{1}{c^2} \cdot \frac{(u_1 - v)^2}{\left(1 - \frac{u_1 v}{c^2}\right)^2} = \frac{c^2\left(1 - \frac{u_1 v}{c^2}\right)^2 - (u_1 - v)^2}{c^2\left(1 - \frac{u_1 v}{c^2}\right)^2},$$

so that

$$\begin{aligned}\left(1 - \frac{u_1'^2}{c^2}\right)\left(1 - \frac{u_1 v}{c^2}\right)^2 &= \frac{1}{c^2}\left\{\left(c^2 - 2u_1 v + \frac{u_1^2 v^2}{c^2}\right) - (u_1^2 - 2u_1 v + v^2)\right\} \\ &= 1 - \frac{1}{c^2}(u_1^2 + v^2) + \frac{u_1^2 v^2}{c^4}.\end{aligned} \tag{5}$$

Substituting (5) in (4) we get

$$\gamma_1' u_1' = \frac{u_1 - v}{\sqrt{1 - \frac{1}{c^2}(u_1^2 + v^2) + \frac{u_1^2 v^2}{c^4}}}. \tag{6}$$

Comparing (6) with (3), we obtain

$$\gamma_1' u_1' = \gamma\gamma_1\,(u_1 - v). \tag{7}$$

Now suppose that there are a number of such particles moving in a straight line, subject to the condition that both the mass and the momentum are invariant as measured by S, so that we have

$$\Sigma m_1 = \text{constant}, \qquad \Sigma m_1 u_1 = \text{constant}. \tag{8}$$

Since γ and v are also constants, being the same for all the particles, therefore from (8) we get

$$\Sigma m_1\gamma v = \text{constant}, \quad \Sigma m_1 u_1\gamma = \text{constant},$$

or subtracting the first equation from the second,

$$\Sigma m_1\gamma\,(u_1 - v) = \text{constant}, \tag{9}$$

where the summation extends over all the particles.

But from (7) we have

$$\gamma(u_1 - v) = \frac{\gamma_1{}' u_1{}'}{\gamma_1}.$$

Substituting this in (9), we obtain

$$\Sigma m_1 \frac{\gamma_1{}' u_1{}'}{\gamma_1} = \text{constant}. \tag{10}$$

On the other hand, according to our fundamental hypothesis, momentum must also be conserved for the observer S′, so that

$$\Sigma m_1{}' u_1{}' = \text{constant}. \tag{11}$$

The results (10) and (11) will agree if

$$m_1 \frac{\gamma_1{}'}{\gamma_1} = m_1{}' \text{ or if } \frac{m_1}{\gamma_1} = \frac{m_1{}'}{\gamma_1{}'}. \tag{12}$$

Similarly for all other particles. Thus

$$\frac{m_1}{\gamma_1} = \frac{m_1{}'}{\gamma_1{}'} = \text{an absolute constant}$$
$$= m_0 \text{ (say)},$$

so that for every particle, and for any frame of reference, if u is the velocity of the particle relative to that frame, then

$$m = m_0\gamma = \frac{m_0}{\sqrt{1 - \dfrac{u^2}{c^2}}}, \tag{13}$$

showing that the mass of a particle is not constant but increases with the velocity. This increase of mass has been verified experimentally for the cathode rays, and the radioactive β-rays.

For a particle at rest, $u = 0$, and therefore $m = m_0$. m_0 is called therefore the "rest-mass," and the Newtonian mechanics is characterised by the fact that it always uses the invariant rest-mass m_0 instead of the actual variable mass m. For u very small compared to c, this gives a good approximation, but for velocities comparable to that of light, as occur in the atomic theory, Newtonian mechanics is no longer applicable.

1·6 (2). *Relativistic Equations of Motion and Energy.*

Newton's equations of motion (2) § 1·6 are taken over in the relativity theory, with this difference, however, that the mass m of the particle is no longer constant, but changes with its velocity according to the law (13) § 1·6 (1). The components of momentum are therefore

$$p_x = \frac{m_0\dot{x}}{\sqrt{1-\beta^2}}, \quad p_y = \frac{m_0\dot{y}}{\sqrt{1-\beta^2}}, \quad p_z = \frac{m_0\dot{z}}{\sqrt{1-\beta^2}}, \tag{1}$$

where we have written

$$\beta = \frac{u}{c} = \frac{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}{c}. \tag{2}$$

The relativistic equations of motions are therefore

$$\frac{dp_x}{dt} = F_x, \quad \frac{dp_y}{dt} = F_y, \quad \frac{dp_z}{dt} = F_z,$$

or

$$\frac{d}{dt}\frac{m_0\dot{x}}{\sqrt{1-\beta^2}} = F_x, \text{ etc.} \tag{3}$$

Multiplying these equations by $\dot{x}$, $\dot{y}$, $\dot{z}$ respectively, and adding we get corresponding to (3) of § 1·6:

$$m_0\left\{\dot{x}\frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{1-\beta^2}}\right)+\dot{y}\frac{d}{dt}\left(\frac{\dot{y}}{\sqrt{1-\beta^2}}\right)+\dot{z}\frac{d}{dt}\left(\frac{\dot{z}}{\sqrt{1-\beta^2}}\right)\right\}$$
$$= F_x\dot{x} + F_y\dot{y} + F_z\dot{z},$$

or

$$m_0\left\{\frac{\dot{x}x+\dot{y}\dot{y}+\dot{z}z}{\sqrt{1-\beta^2}} - \frac{\beta\dot{\beta}}{(1-\beta^2)^{\frac{3}{2}}}(\dot{x}^2+\dot{y}^2+\dot{z}^2)\right\}$$
$$= F_x\dot{x} + F_y\dot{y} + F_z\dot{z}. \tag{4}$$

But from (2) we have

$$\beta^2 = \frac{1}{c^2}(\dot{x}^2+\dot{y}^2+\dot{z}^2),$$

so that

$$\beta\dot{\beta} = \frac{1}{c^2}(\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z}).$$

Substituting this in (4) and remembering that $\dot{x}^2+\dot{y}^2+\dot{z}^2 = c^2u^2$, we get

$$m_0\left\{\frac{\beta\dot{\beta}c^2}{\sqrt{1-\beta^2}}+\frac{\beta\dot{\beta}c^2\beta^2}{(1-\beta^2)\sqrt{1-\beta^2}}\right\} = F_x\dot{x} + F_y\dot{y}\ F_z\dot{z},$$

or

$$m_0\frac{\beta\dot{\beta}c^2}{(1-\beta^2)}\frac{\{(1-\beta^2)+\beta^2\}}{\sqrt{1-\beta^2}} = F_x\dot{x} + F_y\dot{y} + F_z\dot{z},$$

or

$$m_0 \frac{\beta\dot{\beta}c^2}{(1-\beta^2)^{\frac{3}{2}}} = F_x\dot{x}\ F_y\dot{y}\ F_z\dot{z}. \tag{5}$$

Integrating this equation with respect to time, we get

$$m_0c^2 \int \frac{\beta\dot{\beta}}{(1-\beta^2)^{\frac{3}{2}}}\,dt = \int (F_x\dot{x} + F_y\dot{y} + F_z\dot{z})\,dt,$$

or

$$\begin{aligned} \frac{m_0c^2}{\sqrt{1-\beta^2}} &= \int (F_x dx + F_y dy + F_z dz) + A, \\ &= -V + A, \end{aligned}$$

where V is the potential energy of the external field, and A is a constant. Thus

$$\frac{m_0c^2}{\sqrt{1-\beta^2}} + V = \text{constant}. \tag{6}$$

On the other hand, if we denote the kinetic energy by T as usual, we have the equation of energy

$$T + V = \text{constant}. \tag{7}$$

Comparing (7) with (6) we see that the kinetic energy T must be given by

$$T = \frac{m_0c^2}{\sqrt{1-\beta^2}} + A', \tag{8}$$

where A′ is some constant which must be determined. For this we remark that if u is very small compared to c, then the kinetic energy should be given by Newton's formula $\frac{1}{2}m_0u^2$, so that from (8) we must get

$$\begin{aligned} \tfrac{1}{2}m_0u^2 &= m_0c^2(1-\beta^2)^{-\frac{1}{2}} + A' \\ &= m_0c^2(1 + \tfrac{1}{2}\beta^2) + A' \\ &= m_0c^2 + \tfrac{1}{2}m_0u^2 + A'. \end{aligned}$$

Thus we obtain

$$A' = -m_0c^2, \tag{9}$$

and the kinetic energy is given by (8) and (9):

$$\begin{aligned} T &= \frac{m_0c^2}{\sqrt{1-\beta^2}} - m_0c^2 \\ &= mc^2 - m_0c^2 = c^2(m - m_0). \end{aligned} \tag{10}$$

m_0c^2 is called the " rest-energy " or the " internal energy of constitution " of the material point, which is assumed to be possessed by the particle at rest. This corresponds to the fact that in mechanics only changes of energy are considered, so that the total energy of a system is left vague to the extent of an arbitrary additive constant. We assume, therefore, that this additive constant for a particle of rest mass m_0 is m_0c^2.

Thus the whole energy of a particle moving freely, *i.e.*, in the absence of an external field, is given by

$$\begin{aligned} E &= T + m_0c^2 \\ &= c^2(m - m_0) + m_0c^2, \end{aligned}$$

so that

$$E = mc^2, \quad \text{or } m = \frac{E}{c^2}. \tag{11}$$

This is the well-known formula of Einstein's giving the inertia of energy, and it has been abundantly verified by experiments. It is of fundamental importance in the recent developments of quantum mechanics given in Chapter VI. The formula (11) signifies further that the two fundamental conceptions of mass and energy are identical.

1·7. *Minkowski's Four-Dimensional Vectors.*

In § 1·5, we saw that the external world is a four-dimensional continuum, in which the four co-ordinates of an event are given by

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict. \tag{1}$$

Our equations of motion therefore must have an invariant form in this continuum, and must have four components. Thus corresponding to the position "four-vector" (short for four-dimensional vector, to distinguish it from the ordinary three-dimensional vector in Euclidean space), we must have a velocity "four-vector" (u_r), with components u_1, u_2, u_3, u_4 and a force "four-vector" (F_r) with components F_1, F_2, F_3, F_4. We must naturally find the significance of the components of these four-vectors in terms of the classical quantities.

Now in classical mechanics, the components u_x, u_y, u_z of the velocity are given by differentiating the co-ordinates x, y, z, with respect to the time t:

$$u_x = \frac{dx}{dt}, \quad u_y = \frac{dy}{dt}, \quad u_z = \frac{dz}{dt}. \tag{2}$$

In the classical theory, t is invariant for all systems of co-ordinates.

But in the relativity theory t is not invariant. However, in § 1·5 we have seen that there exists an invariant time, *viz.*, the eigentime τ, connected with t according to the relation

$$d\tau = \sqrt{1 - \frac{u^2}{c^2}}\, dt = \sqrt{1-\beta^2}\, dt, \tag{3}$$

where u is the velocity of the particle. To get the components of the velocity four-vector therefore, we must differentiate the co-ordinates (1) with respect to the invariant eigentime τ. Thus

$$\begin{aligned} u_1 &= \frac{dx_1}{d\tau} = \frac{dx}{dt}\cdot\frac{dt}{d\tau} = \frac{u_x}{\sqrt{1-\beta^2}}, \\ u_2 &= \frac{dx_2}{d\tau} = \frac{dy}{dt}\cdot\frac{dt}{d\tau} = \frac{u_y}{\sqrt{1-\beta^2}}, \\ u_3 &= \frac{dx_3}{d\tau} = \frac{dz}{dt}\cdot\frac{dt}{d\tau} = \frac{u_z}{\sqrt{1-\beta^2}}, \\ u_4 &= \frac{dx_4}{d\tau} = \frac{d}{dt}(ict)\cdot\frac{dt}{d\tau} = \frac{ic}{\sqrt{1-\beta^2}}. \end{aligned} \tag{4}$$

The momentum four-vector $[p]$ is given, in analogy with the classical momentum, by multyplying the velocity four-vector $[u]$ by the invariant mass m_0; so that the components of momentum are

$$\begin{aligned} p_1 &= m_0 u_1 = \frac{m_0 u_x}{\sqrt{1-\beta^2}}, \quad & p_2 &= m_0 u_2 = \frac{m_0 u_y}{\sqrt{1-\beta^2}}, \\ p_3 &= m_0 u_3 = \frac{m_0 u_z}{\sqrt{1-\beta^2}}, \quad & p_4 &= m_0 u_4 = \frac{i m_0 c}{\sqrt{1-\beta^2}}. \end{aligned} \tag{5}$$

From equation (11) § 1·6 (2) and the last equation of (5), we find that the time component p_4 of momentum is given by

$$\begin{aligned} p_4 &= \frac{i m_0 c}{\sqrt{1-\beta^2}} = imc = \frac{imc^2}{c} \\ &= \frac{i\,\mathrm{E}}{c}. \end{aligned} \tag{6}$$

Thus the four-dimensional form of the equations of motion is

$$\frac{dp_r}{d\tau} = \mathrm{F}_r, \quad (r = 1, 2, 3, 4). \tag{7}$$

We have now to find the physical significance of the components F_r of the force four-vector. On re-introducing the ordinary time t from (3), the equations (7) become

$$\frac{dp_r}{dt} = \mathrm{F}_r \sqrt{1-\beta^2},$$

or, since $p_r = m_0 u_r$,

$$m_0 \frac{du_r}{dt} = \mathrm{F}_r \sqrt{1-\beta^2}, \quad (r = 1, 2, 3, 4). \tag{8}$$

Substituting for u_1, u_2, u_3 their values (4), we get .

$$
\begin{aligned}
m_0 \frac{d}{dt}\left(\frac{u_x}{\sqrt{1-\beta^2}}\right) &= F_1 \sqrt{1-\beta^2},\\
m_0 \frac{d}{dt}\left(\frac{u_y}{\sqrt{1-\beta^2}}\right) &= F_2 \sqrt{1-\beta^2} \qquad (9)\\
m_0 \frac{d}{dt}\left(\frac{u_z}{\sqrt{1-\beta^2}}\right) &= F_3 \sqrt{1-\beta^2}\cdot
\end{aligned}
$$

Comparing these equations with the equations (3) § 1·6 (2), we see that the first three components of the force four-vector (F), and the three components of the Newtonian force-vector $\vec{F}$ are connected by the relations

$$
F_1 = \frac{F_x}{\sqrt{1-\beta^2}}, \quad F_2 = \frac{F_y}{\sqrt{1-\beta^2}}, \quad F_3 = \frac{F_z}{\sqrt{1-\beta^2}}\cdot \qquad (10)
$$

In order to find the significance of the fourth component F_4, we multiply the equations (7) with u_r, and sum from 1 to 4, getting, since, $p_r = m_0 u_r$,

$$
\sum_{r=1}^{4} m_0 u_r \frac{du_r}{d\tau} = \sum_{r=1}^{4} F_r u_r,
$$

or

$$
\tfrac{1}{2} m_0 \frac{d}{d\tau}\left(\sum_r u_r^2\right) = \sum_r F_r u_r. \qquad (11)
$$

But on account of (4) we have

$$
\sum_{r=1}^{4} u_r^2 = \frac{u_x^2 + u_y^2 + u_z^2 - c^2}{1-\beta^2} = \frac{u^2 - c^2}{1 - \frac{u^2}{c^2}} = -c^2, \qquad (12)
$$

showing that the velocity four-vector has the constant magnitude *ic*. From (11) we get, therefore,

$$
\sum_{r=1}^{4} F_r u_r = 0,
$$

i.e., on substituting the values of F_1, F_2, F_3 from (10) and the values of u_1, u_2, u_3, u_4 from (4)

$$
\frac{F_x u_x + F_y u_y + F_z u_z}{1-\beta^2} + \frac{F_4 ic}{\sqrt{1-\beta^2}} = 0.
$$

Thus, we get the required value of F_4

$$
F_4 = \frac{i}{c\sqrt{1-\beta^2}}\left(F_x u_x + F_y u_y + F_z u_z\right). \qquad (13)
$$

Now the fourth component of the equation of motion (8) is

$$m_0 \frac{du_4}{dt} = F_4 \sqrt{1-\beta^2},$$

which becomes on writing the value of u_4 from (4) and of F_4 from (13):

$$m_0 \frac{d}{dt}\left(\frac{ic}{\sqrt{1-\beta^2}}\right) = \frac{i}{c}\left(F_x u_x + F_y u_y + F_z u_z\right),$$

or

$$\frac{d}{dt}\left(\frac{m_0 c^2}{\sqrt{1-\beta^2}}\right) = F_x u_x + F_y u_y + F_z u_z. \qquad (14)$$

If we remark that $\frac{m_0 c^2}{\sqrt{1-\beta^2}}$ is just the energy E of the body as given in (11) § 1·6 (2), and that the right-hand side of (14) is the work done by the external forces in unit time, then we see that (14) is only the expression of the energy theorem.

CHAPTER II.

GENERAL THEORY OF DYNAMICS.

2·1. *Hamilton's Variation Principle.*

SUPPOSE that (x, y, z) are the Cartesian co-ordinates of a particle, and that it is acted upon by a force whose components are X, Y, Z. Let the components of velocity be denoted by $u = \frac{dx}{dt} = \dot{x}$, $v = \dot{y}$, $w = \dot{z}$, and those of momentum by $p_x = mu$, $p_y = mv$, $p_z = mw$. The kinetic energy of the particle is

$$T = \tfrac{1}{2} m (u^2 + v^2 + w^2),$$

so that

$$p_x = \frac{\partial T}{\partial u} = \frac{\partial T}{\partial \dot{x}}, \quad p_y = \frac{\partial T}{\partial v} = \frac{\partial T}{\partial \dot{y}}, \quad p_z = \frac{\partial T}{\partial w} = \frac{\partial T}{\partial \dot{z}}. \tag{1}$$

Newton's equations of motion are

$$m\dot{u} = X, \quad m\dot{v} = Y, \quad m\dot{w} = Z, \tag{2}$$

or,

$$\dot{p}_x = X, \quad \dot{p}_y = Y, \quad \dot{p}_z = Z. \tag{3}$$

If the external forces are conservative, so that there exists a potential energy V, we have

$$X = -\frac{\partial V}{\partial x}, \quad Y = -\frac{\partial V}{\partial y}, \quad Z = -\frac{\partial V}{\partial z}. \tag{4}$$

The equations of motion now become

$$m\dot{u} = -\frac{\partial V}{\partial x}, \quad m\dot{v} = -\frac{\partial V}{\partial y}, \quad m\dot{w} = -\frac{\partial V}{\partial z}, \tag{5}$$

Multiplying these equations by u, v, w respectively and adding, we get

$$\begin{aligned} m(u\dot{u} + v\dot{v} + w\dot{w}) &= -\left(\frac{\partial V}{\partial x} u + \frac{\partial V}{\partial y} v + \frac{\partial V}{\partial z} w\right) \\ &= -\left(\frac{\partial V}{\partial x} \dot{x} + \frac{\partial V}{\partial y} \dot{y} + \frac{\partial V}{\partial z} \dot{z}\right) \\ &= -\frac{dV}{dt}. \end{aligned} \tag{6}$$

Integrating (6) we find

$$\tfrac{1}{2} m (u^2 + v^2 + w^2) + V = \text{constant}, \tag{7}$$

This is the law of conservation of mechanical energy for forces such as gravitational, electric and magnetic, which have a potential, and are said to be conservative.

Let us now compare this motion with a slightly different motion in which Newton's laws are not obeyed. In this motion, let the co-ordinates of m at the instant t be (x', y', z'), and let the components of velocity at the same instant be u', v', w', so that $u' = \frac{dx'}{dt}$, etc. The co-ordinates of m in the actual motion at the same instant t are (x, y, z) as above, and we assume that the modified motion differs so little from the actual that $x' - x$, $u' - u$, etc., are very small. This means that if we write

$$x' - x = \delta x, \quad u' - u = \delta u, \text{etc.},$$

we can neglect higher powers of δx, δu, etc.

If we multiply the equations (5) by δx, δy, δz respectively and add, we get

$$m\,(\dot{u}\delta x + \dot{v}\,\delta y + \dot{w}\delta z) = -\left(\frac{\partial V}{\partial x}\,\delta x + \frac{\partial V}{\partial y}\,\delta y + \frac{\partial V}{\partial z}\,\delta z\right). \tag{8}$$

Now,

$$\frac{du}{dt}\,\delta x = \frac{d}{dt}\,(u\delta x) - u\,\frac{d}{dt}\,(\delta x), \tag{9}$$

and

$$\frac{d}{dt}\,(\delta x) = \frac{d}{dt}\,(x' - x) = u' - u = \delta u$$
$$= \delta\left(\frac{dx}{dt}\right). \tag{10}$$

From this last equation we see that the operations of differentiation $\frac{d}{dt}$ and of variation δ are commutative, and we can change their order without affecting the result.

Substituting (10) in (9) we get:

$$\frac{du}{dt}\,\delta x = \frac{d}{dt}\,(u\delta x) - u\delta u. \tag{11}$$

We get similar expressions for v and w. Substituting these in (8) we have

$$m\left\{\frac{d}{dt}\,(u\delta x + v\delta y + w\delta z) - (u\delta u + v\delta v + w\delta w)\right\}$$
$$= -\left(\frac{\partial V}{\partial x}\,\delta x + \frac{\partial V}{\partial y}\,\delta y + \frac{\partial V}{\partial z}\,\delta z\right) = -\,\delta V. \tag{12}$$

If the dynamical system consists of a number of particles, an equation of this kind is true for each one of them, and at every instant of the motion. It is, moreover true whatever the displaced motion may be, provided only that it differs very little from the actual motion. Thus for the whole system

$$\Sigma m \left\{\frac{d}{dt}(u\delta x + v\delta y + w\delta z) - (u\delta u + v\delta v + w\delta w)\right\} = -\ \delta V. \tag{13}$$

In this equation V denotes the potential energy for the whole system. Now,

$$T = \Sigma \tfrac{1}{2} m (u^2 + v^2 + w^2), \quad T' = \Sigma \frac{m}{2} (u'^2 + v'^2 + w'^2),$$

therefore

$$\begin{aligned} \delta T = T' - T &= \tfrac{1}{2} \Sigma m \{(u'^2 - u^2) + (v'^2 - v^2) + (w'^2 - w^2)\} \\ &= \tfrac{1}{2} \Sigma m \{(u + \delta u)^2 - u^2 + (v + \delta v)^2 - v^2 + (w + \delta w)^2 - w^2\} \\ &= \Sigma m (u\delta u + v\delta v + w\delta w), \end{aligned} \tag{14}$$

neglecting $(\delta u)^2$, etc.

Substituting this in (13), we get

$$\Sigma m \frac{d}{dt} (u\delta x + v\delta y + w\delta z) = \delta(T - V). \tag{15}$$

We write

$$L = T - V. \tag{16}$$

L is called the Lagrangian function. Equation (15) is true at every instant of the motion. We integrate it for any two instants of the motion, say from $t = t_1$ to $t = t_2$:

$$\int_{t_1}^{t_2} \left\{\Sigma m \frac{d}{dt} (u\delta x + v\delta y + w\delta z)\right\} dt = \int_{t_1}^{t_2} \delta L\, dt.$$

or

$$\Big[\Sigma m (u\delta x + v\delta y + w\delta z)\Big]_{t_1}^{t_2} = \int_{t_1}^{t_2} \delta L\, dt. \tag{17}$$

The displaced motion has been so far quite arbitrary, except for the fact that it was supposed to differ very little from the actual motion. We now impose the further restriction that at times t_1 and t_2 the configurations in the displaced motion are the same as those in the actual motion. Then we have at instants t_1 and t_2

$$\delta x = \delta y = \delta z = 0$$

for all particles. So that

$$\Big[\Sigma m (u\delta x + v\delta y + w\delta z)\Big]_{t_1}^{t_2} = 0.$$

Therefore, from (17) we have

$$\int_{t_1}^{t_2} \delta L dt = 0. \tag{18}$$

But

$$\int_{t_1}^{t_2} \delta L\, dt = \int_{t_1}^{t_2} (L' - L)\, dt = \int_{t_1}^{t_2} L'dt - \int_{t_1}^{t_2} Ldt$$

$$= \delta \int_{t_1}^{t_2} Ldt,$$

so that equation (18) becomes

$$\delta \int_{t_1}^{t_2} Ldt = 0. \tag{19}$$

This is Hamilton's variation principle. It depends only on the kinetic and potential energies of the system and is independent of any co-ordinates. It can be shown that conversely from this single equation, we can determine the motion of all the known parts of the system as soon as the kinetic and potential energies are known.

Now if we denote the constant total energy of the system by E, so that $E = T + V$, then

$$L = T - V = T - (E - T) = 2\,T - E, \tag{20}$$

and the equation (19) becomes

$$0 = \delta \int_{t_1}^{t_2} Ldt = \delta \int_{t_1}^{t_2} (2T - E)\, dt.$$

But since E is constant $\delta \int_{t_1}^{t_2} Edt = 0$, so that

$$\delta \int_{t_1}^{t_2} 2Tdt = 0. \tag{21}$$

The function $S = \int_{t_1}^{t_2} 2Tdt$ is called the action of the motion, and (21) shows that the action is always an extremum. As a matter of fact, for an actual motion, the action is always a minimum, and that is why Hamilton's variation principle is also called the "principle of least action", first enunciated by Maupertius in 1744,

2 · 2. *Lagrange's Equations.*

Suppose we have a dynamical system consisting of material particles and rigid bodies, which can move either freely or connected in any manner. We need a finite number, say n, of quantities to fix a configuration of this system at any time. For instance, if a rigid body is rotating about an axis fixed in itself, its configuration at any time will be determined if we know a single angle θ. Similarly, if a particle is moving on a spherical surface, its position will be fixed by the two angles θ, ϕ. Thus in the first example, the Cartesian co-ordinates x, y, z of any point of the body will be functions of θ only, whereas in the second system, the Cartesian co-ordinates of the particle will be functions of the two angles θ and ϕ. Of course, time is the independent variable in both systems, on which θ and ϕ depend.

The variables θ, ϕ are called "generalised co-ordinates," and n is called the "degree of freedom". Usually we denote the generalised co-ordinates by $q_1, q_2, \cdots, q_n$, each of them being a function of t. Then the Cartesian co-ordinates x, y, z of any particle of the system are functions of the q's,

$$x = x(q_1, \cdots, q_n), \quad y = y(q_1, \cdots, q_n), \quad z = z(q_1, \cdots, q_n),$$

so that

$$\begin{aligned}
\frac{dx}{dt} &= \frac{\partial x}{\partial q_1}\dot{q}_1 + \frac{\partial x}{\partial q_2}\dot{q}_2 + \cdots + \frac{\partial x}{\partial q_n}\dot{q}_n,\\
\frac{dy}{dt} &= \frac{\partial y}{\partial q_1}\dot{q}_1 + \quad \cdots \quad + \frac{\partial y}{\partial q_n}\dot{q}_n,\\
\frac{dz}{dt} &= \frac{\partial z}{\partial q_1}\dot{q}_1 + \quad \cdots \quad + \frac{\partial z}{\partial q_n}\dot{q}_n.
\end{aligned}$$

Thus, we know the velocity of each particle if we know x, y, z in terms of q's and also if we know $\dot{q}_1, \cdots, \dot{q}_n$. These last quantities are called the generalised components of velocities.

The kinetic energy $T = \frac{1}{2}\Sigma m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$ is now seen to be a quadratic function of $\dot{q}_1, \cdots, \dot{q}_n$, the coefficients being functions of $q_1, \cdots, q_n$. The potential energy V depends only on the configuration of the system, and is a function of $q_1, \cdots, q_n$ only.

Thus the function $L = T - V$ is a function of $q_1, \cdots, q_n, \dot{q}_1, \cdot\cdot \dot{q}_n$:

$$L = L(q_1, \cdots q_n ; \dot{q}_1, \cdots, \dot{q}_n).$$

The corresponding function L′ in the displaced motion is the same function of $q_1 + \delta q_1, \cdots, \dot{q}_n + \delta \dot{q}_n$:

$$L' = L'(q_1 + \delta q_1, \cdots q_n + \delta q_n;\ \dot{q}_1 + \delta \dot{q}_1, \cdots, \dot{q}_n + \delta \dot{q}_n),$$

so that from Taylor's theorem we have:

$$\delta L = L' - L = \frac{\partial L}{\partial q_1}\delta q_1 + \cdots + \frac{\partial L}{\partial q_n}\delta q_n + \frac{\partial L}{\partial \dot{q}_1}\delta \dot{q}_1 + \cdots + \frac{\partial L}{\partial \dot{q}_n}\delta \dot{q}_n.$$

Therefore from Hamilton's theorem we get

$$0 = \int_{t_1}^{t_2} \delta L dt = \int_{t_1}^{t_2} \left\{ \sum_{r=1}^{n} \frac{\partial L}{\partial q_r}\delta q_r + \sum_{r=1}^{n} \frac{\partial L}{\partial \dot{q}_r}\delta \dot{q}_r \right\} dt. \qquad (1)$$

But

$$\delta \dot{q}_r = \delta\left(\frac{dq_r}{dt}\right) = \dot{q}_r' - q_r' = \frac{d}{dt}(q_r' - q_r) = \frac{d}{dt}(\delta q_r).$$

Therefore

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_r}\delta \dot{q}_r dt = \int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_r}\frac{d}{dt}(\delta q_r) dt$$

$$= \left[\frac{\partial L}{\partial \dot{q}_r}\delta q_r\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_r}\right)\delta q dt.$$

If we take the displaced motion such that at the times t_1 and t_2 the q's have the same value, then $\delta q_r = 0$ both at t_1 and t_2. Thus we have

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_r}\delta \dot{q}_r dt = -\int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_r}\right)\delta q_r dt. \qquad (2)$$

Let us write

$$p_r = \frac{\partial T}{\partial \dot{q}_r} = \frac{\partial (T - V)}{\partial \dot{q}_r} = \frac{\partial L}{\partial \dot{q}_r}, \quad (r = 1, 2, \cdots n) \qquad (3)$$

since V does not depend on $\dot{q}_r$. p_r is called the generalised momentum conjugate to the co-ordinate q_r. Then from equations (1), (2) and (3) we get:

$$\int_{t_1}^{t_2} \delta L dt = \int_{t_1}^{t_2} \left\{ \sum_{r=1}^{n} \left(\frac{\partial L}{\partial q_r} - \frac{dp_r}{dt}\right)\delta q_r \right\} dt = 0. \qquad (4)$$

Since t_1 and t_2 are entirely arbitrary, we have

$$\sum_{r=1}^{n} \left(\frac{\partial L}{\partial q_r} - \frac{dp_r}{dt}\right)\delta q_r = 0. \qquad (5)$$

But all the δq's are independent, therefore (5) can be true only when each of the coefficients separately vanishes, so that

$$\frac{dp_r}{dt} - \frac{\partial L}{\partial q_r} = 0, \quad (r = 1, 2, \cdots, n), \qquad (6)$$

or

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_r}\right) - \frac{\partial T}{\partial q_r} = - \frac{\partial V}{\partial q_r}, \quad (r = 1, 2, \cdots n). \tag{7}$$

These are n differential equations of the second order, and they enable us to determine $q_1, \cdots, q_n$ as functions of time. They are called " Lagrange's Equations ".

Supposing that the forces, and therefore the potential function V, depend on time as well as on co-ordinates, then the Lagrangian L also depends on t, so that

$$L = L\,(q_1, \cdots, q_n, \dot{q}_1 \cdots, \dot{q}_n, t). \tag{8}$$

It is easily seen that Hamilton's Principle and Lagrange's Equations are still true, but the mechanical energy is no longer conserved.

2·3. *Hamilton's Canonical Equations.*

In the last section, we saw that the motion of a dynamical system is determined by a system of n differential equations of the second order, where n is the number of degrees of freedom of the system. We shall see now that the motion can also be determined by a system of $2n$ differential equations of the first order which can be easily derived from Lagrange's Equations.

We introduce a function T′ defined by

$$T' = \sum_{r=1}^{n} p_r \dot{q}_r - T; \tag{1}$$

then

$$dT' = \sum_{r=1}^{n} p_r d\dot{q}_r + \sum_{r=1}^{n} \dot{q}_r dp_r - \sum_{r=1}^{n} \frac{\partial T}{\partial q_r} dq_r - \sum_{r=1}^{n} \frac{\partial T}{\partial \dot{q}_r} d\dot{q}_r.$$

But from the definition of p_r we have $p_r = \dfrac{\partial T}{\partial \dot{q}_r}$, so that the first and the last sums on the right-hand side cancel out, and we are left with

$$dT' = \sum_{r=1}^{n} \dot{q}_r dp_r - \sum_{r=1}^{n} \frac{\partial T}{\partial q_r} dq_r, \tag{2}$$

which shows that T′ is a function of p's and q's only.

Now in Newton's mechanics, T is a homogeneous function of the second degree in $\dot{q}$'s, so that from Euler's well-known theorem in Calculus, we have

$$\sum_{r=1}^{n} \frac{\partial T}{\partial \dot{q}_r} \dot{q}_r = 2T.$$

therefore, since $p_r = \dfrac{\partial T}{\partial \dot{q}_r}$,

$$\sum_{n=1}^{n} p_r \dot{q}_r = 2T\,; \tag{3}$$

so that from (1) and (3) we find

$$T' = \sum_{r=1}^{n} p_r \dot{q}_r - T = 2T - T = T. \tag{4}$$

Equation (4) expresses the fact that T′ represents the kinetic energy, but whereas T represents the kinetic energy in the variables q_r, $\dot{q}_r$, *i.e.*, in co-ordinates and velocities, T′ represents the kinetic energy in q_r and p_r, *i.e.*, in co-ordinates and momenta.

We define another function H as

$$H = T' + V = \sum_{r=1}^{n} p_r \dot{q}_r - T + V = \sum_{r=1}^{n} p_r \dot{q}_r - L\,; \tag{5}$$

then we see that H is the total energy expressed in the co-ordinates q and the momenta p, so that we can write $H = H(p, q)$ where p stands for all the $p_1, \cdots p_n$ and q for all the $q_1, \cdots q_n$. This function $H(p, q)$ is called the "Hamiltonian" of the dynamical system, and we remark again that it represents the total energy (kinetic energy + potential energy) expressed as a function of co-ordinates and momenta.

From the equation (2) we see that

$$\frac{\partial T'}{\partial q_r} = -\frac{\partial T}{\partial q_r}\,; \quad \frac{\partial T'}{\partial p_r} = \dot{q}_r, \quad (r = 1, \cdots, n). \tag{6}$$

Thus Lagrange's Equations (6) § 2 · 2 may be written

$$\frac{dp_r}{dt} = \frac{\partial L}{\partial q_r} = \frac{\partial}{\partial q_r}(T - V) = -\frac{\partial}{\partial q_r}(T' + V) = -\frac{\partial H}{\partial q_r}. \tag{7}$$

Also from the second part of (6) we have, since $\dfrac{\partial V}{\partial p_r} = 0$,

$$\frac{dq_r}{dt} = \frac{\partial T'}{\partial p_r} = \frac{\partial}{\partial p_r}(T' + V) = \frac{\partial H}{\partial p_r}. \tag{8}$$

Thus we have the system of $2n$ equations

$$\left.\begin{aligned} \dot{q}_r \equiv \frac{dq_r}{dt} &= \frac{\partial H}{\partial p_r}, \quad (r = 1, 2, \cdots, n), \\ \dot{p}_r \equiv \frac{dp_r}{dt} &= -\frac{\partial H}{\partial q_r}, \quad (r = 1, 2, \cdots, n). \end{aligned}\right\} \tag{9}$$

These are called "Hamilton's Canonical Equations" and they depend only on the total energy. With their help we can determine the motion of the various parts of the system even if we do not know the actual mechanism such as the interacting forces, etc. They represent the simplest form in which the generalised dynamical equations can be expressed, and they form the basis of various investigations in higher dynamics, astronomy and physics.

2 · 4. *Relativistic Lagrangian and Hamiltonian.*

In the previous sections we saw that the Lagrangian function L is given by

$$L = T - V, \tag{1}$$

where the kinetic energy T is expressed as a function of p_r, q_r, by means of the relations

$$p_r = \frac{\partial T}{\partial \dot{q}_r} = \frac{\partial L}{\partial \dot{q}_r}. \tag{2}$$

We have also seen that Newton's equations of motion are equivalent to Lagrange's equations

$$\frac{dp_r}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_r}\right) = \frac{\partial L}{\partial q_r}. \tag{3}$$

Now the relativistic kinetic energy of a particle of rest mass m_0, moving with the velocity $u = \beta c$, is

$$T = m_0 c^2 \left(\frac{1}{\sqrt{1 - \beta^2}} - 1\right). \tag{4}$$

We wish to express the relativity equations of motion in the Lagrangian form (3). But we find that $\dfrac{\partial T}{\partial \dot{x}}$ does not give the correct relativistic momentum $p_x = \dfrac{m_0 \dot{x}}{\sqrt{1 - \beta^2}}$. We must therefore find another function T^* such that

$$\frac{m_0 \dot{x}}{\sqrt{1 - \beta^2}} = p_x = \frac{\partial T^*}{\partial \dot{x}}; \tag{5}$$

similarly for the y- and z-components. Then we have:

$$dT^* = \frac{\partial T^*}{\partial \dot{x}} d\dot{x} + \frac{\partial T^*}{\partial \dot{y}} d\dot{y} + \frac{\partial T^*}{\partial \dot{z}} d\dot{z}$$

$$= \frac{m_0}{\sqrt{1 - \beta^2}} (\dot{x} d\dot{x} + \dot{y} d\dot{y} + \dot{z} d\dot{z}) = \frac{m_0 c^2 \beta d\beta}{\sqrt{1 - \beta^2}},$$

because $c^2\beta^2 = u^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$. Integrating, we get:

$$T^* = -m_0 c^2 \sqrt{1 - \beta^2} + A,$$

where A is a constant of integration. Now for a velocity u which is very small compared to c, the function T^* must reduce to the classical kinetic energy $\frac{1}{2} m_0 u^2$. Therefore

$$\tfrac{1}{2}m_0u^2 = -m_0c^2(1-\tfrac{1}{2}\beta^2) + A,$$
$$= -m_0c^2 + \tfrac{1}{2}m_0u^2 + A,$$

so that

$$A = m_0c^2.$$

Thus the function T^* is determined to be

$$T^* = m_0c^2(1-\sqrt{1-\beta^2}). \tag{6}$$

The relativistic Lagrangian L can now be defined as

$$L = T^* - V = m_0c^2(1-\sqrt{1-\beta^2}) - V, \tag{7}$$

and then the relativistic Lagrangian equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}, \text{ etc.,} \tag{8}$$

are identical with the equations of motion.

As in Newtonian mechanics, we define the relativistic Hamiltonian to be

$$H = p_x\dot{x} + p_y\dot{y} + p_z\dot{z} - L$$
$$= \Sigma\frac{\partial L}{\partial \dot{x}}\dot{x} - L = \Sigma\frac{\partial T^*}{\partial \dot{x}}\dot{x} - L. \tag{9}$$

It can be shown that H still denotes the total energy, *i.e.*, the sum of the kinetic and potential energies.

We have in fact, on account of (5):

$$\Sigma p_x\dot{x} = \Sigma\frac{\partial T^*}{\partial \dot{x}}\dot{x} = \Sigma\frac{m_0\dot{x}}{\sqrt{1-\beta^2}}\dot{x}$$
$$= \frac{m_0}{\sqrt{1-\beta^2}}(\dot{x}^2+\dot{y}^2+\dot{z}^2) = \frac{m_0u^2}{\sqrt{1-\beta^2}} = \frac{m_0\beta^2c^2}{\sqrt{1-\beta^2}}.$$

Therefore, from (9) and (7) and (4):

$$H = \Sigma p_x\dot{x} - L = \frac{m_0\beta^2c^2}{\sqrt{1-\beta^2}} - m_0c^2(1-\sqrt{1-\beta^2}) + V$$
$$= m_0c^2\left(\frac{1}{\sqrt{1-\beta^2}} - 1\right) + V = T + V, \tag{10}$$

showing that H is the total energy.

Now from $p_x = \dfrac{m_0\dot{x}}{\sqrt{1-\beta^2}}$, $p_y = \dfrac{m_0\dot{y}}{\sqrt{1-\beta^2}}$, $p_z = \dfrac{m_0\dot{z}}{\sqrt{1-\beta^2}}$,

we have:

$$p_x^2+p_y^2+p_z^2 = \frac{m_0^2}{1-\beta^2}(\dot{x}^2+\dot{y}^2+\dot{z}^2) = \frac{m_0^2u^2}{1-\beta^2} = \frac{m_0^2c^2\beta^2}{1-\beta^2}.$$

Thus

$$\frac{1}{m_0^2c^2}(p_x^2+p_y^2+p_z^2) = \frac{\beta^2}{1-\beta^2} = \frac{1}{1-\beta^2} - 1.$$

Therefore

$$\frac{1}{1-\beta^2} = \frac{1}{m_0^2c^2}(p_x^2+p_y^2+p_z^2)+1 = \frac{1}{m_0^2c^2}(m_0^2c^2+p_x^2+p_y^2+p_z^2)$$

we find finally from (10):

$$H = c\sqrt{m_0^2c^2+p_x^2+p_y^2+p_z^2} - m_0c^2 + V. \tag{11}$$

With this value of H, the canonical equations of Hamilton:

$$\frac{dx}{dt} = \frac{\partial H}{\partial p_x}, \quad \frac{dp_x}{dt} = -\frac{\partial H}{\partial x}, \tag{12}$$

are verified to be identical with the equation of motion.

2·5. *Canonical or Contact Transformations, Cyclic Variables; the Hamilton-Jacobi Partial Differential Equation.*

Hamilton's canonical equations will determine q_r, p_r $(r = 1, \ldots, n)$ as functions of t, and of $2n$ constants of integration. These $2n$ constants can be chosen at will. For example, we can take the initial co-ordiantes and initial momenta.

Now consider the integral

$$U = \int_0^t L dt = \int_0^t \left(\sum_{r=1}^{n} p_r\dot{q}_r - H\right) dt, \tag{1}$$

then

$$dU = p_1dq_1 + \cdots + p_ndq_n - Hdt, \tag{2}$$

so that

$$\frac{\partial U}{\partial q_1} = p_1, \ \ldots, \ \frac{\partial U}{\partial q_n} = p_n, \ \frac{\partial U}{\partial t} = -H. \tag{3}$$

From the last equation of (3), we get on substituting for p_r in $H(p_1, \cdots, p_n, q_1 \cdots, q_n, t)$:

$$\frac{\partial U}{\partial t} + H\left(q_1, \cdots, q_n, \frac{\partial U}{\partial q_1}, \cdots, \frac{\partial U}{\partial q_n}, t\right) = 0. \tag{4}$$

The function U defined in (1) satisfies the partial differential equation (4) which was given by Jacobi.

If the Hamiltonian H does not depend explicitly on t, we have seen that the total energy is conserved, and

$$H(q_1, \ldots, q_n, p_1, \ldots, p_n) = E \tag{5}$$

The function U then becomes:

$$U = \int_0^t \left(\sum_r p_r\dot{q}_r - E\right) dt = S - Et, \tag{6}$$

where

$$S = \int_0^t \left(\sum_r p_r\dot{q}_r\right) dt = \int_0^t \sum_r p_r dq_r. \tag{7}$$

From (7) we get

$$\frac{\partial S}{\partial q_r} = p_r \ (r = 1, \cdots, n), \tag{8}$$

and then the equation (5) becomes:

$$H\left(q_1, \cdots, q_n, \frac{\partial S}{\partial q_1}, \cdots, \frac{\partial S}{\partial q_n}\right) = E. \tag{9}$$

Now suppose that the Hamiltonian H does not contain explicitly one of the q's, say q_k. Then from Hamilton's canonical equations:

$$\dot{p}_k = \frac{dp_k}{dt} = -\frac{\partial H}{\partial q_k} = 0, \text{ so that } p_k = a_k, \tag{10}$$

where a_x is a constant independent of time. Such a co-ordinate q_k is called a cyclic variable, and p_k is an integral of the equations of motion.

However, if all the q_r are absent from H, we have n integrals of motion provided by the n momenta:

$$p_1 = a_1, \ldots, p_n = a_n. \tag{11}$$

If further, the Hamiltonian does not contain the time explicitly, H is constant, and therefore from the other set of canonical equations, we get

$$\dot{q}_r = \frac{\partial H}{\partial p_r} = \text{constant} = \omega_r \text{ (say)}, \tag{12}$$

so that

$$q_r = \omega_r t + \beta_r, \quad (r = 1, \ldots, n). \tag{13}$$

In (11) and (13) we possess, therefore, $2n$ integrals of motion containing the $2n$ constants a_r, β_r.

We remark that the problem of integrating the canonical equations of motion is reduced to the following problem: "Replace the variables p_r, q_r by new variables P_r, Q_r, such that they have the following two properties:

(1) they preserve the canonical form of the equations of motion;

(2) they transform the Hamiltonian into one which does not contain t, $Q_1, \ldots, Q_n$ explicitly."

Suppose that such new variables P_r, Q_r are obtained by means of the transformation:

$$\begin{aligned} q_r &= q_r(Q_1, Q_2, \ldots, Q_n, P_1, \ldots, P_n, t), \\ p_r &= p_r(Q_1, \ldots, Q_n, P_1, \ldots P_n, t). \end{aligned} \tag{14}$$

We shall determine the condition that this transformation leaves the canonical form of the equations of motion unaltered, that is to say, that the transformation is a "canonical" or a "contact" transformation.

Now we have seen that the canonical equations of motion are equivalent to Hamilton's variation principle

$$\delta \int_0^t (\sum_r p_r \dot{q}_r - \mathrm{H})\, dt = \delta \int_0^t \mathrm{L} dt = 0. \tag{15}$$

If we change to new variables Q_r, P_r, the condition that they are canonically conjugate, *i.e.*, that they satisfy the canonical equations

$$\frac{d\mathrm{Q}_r}{dt} = \frac{\partial \mathrm{K}}{\partial \mathrm{P}_r}, \quad \frac{d\mathrm{P}_r}{dt} = -\frac{\partial \mathrm{K}}{\partial \mathrm{Q}_r}, \tag{16}$$

where K is the new Hamiltonian

$$\mathrm{K}(\mathrm{Q}_r, \mathrm{P}_r, t) = \mathrm{H}(q_r, p_r, t), \tag{17}$$

would be

$$\delta \int_0^t (\sum_r \mathrm{P}_r \dot{\mathrm{Q}}_r - \mathrm{K})\, dt = 0. \tag{18}$$

From Calculus we know that for each one of these equations to be a consequence of the other, it is necessary and sufficient that

$$(\sum_r p_r \dot{q}_r - \mathrm{H}) - (\sum_r \mathrm{P}_r \dot{\mathrm{Q}}_r - \mathrm{K}) = \frac{df}{dt}, \tag{19}$$

where f is any function of t and the canonical variables.

That the condition (19) is sufficient can be seen immediately, for

$$\delta \int_0^t \{(\sum_r p_r \dot{q}_r - \mathrm{H}) - (\sum_r \mathrm{P}_r \dot{\mathrm{Q}}_r - \mathrm{K})\}\, dt = \delta \int_0^t \frac{df}{dt}\, dt = \delta[f]_0^t = 0 \tag{20}$$

since the variables do not alter at the end points.

It can be verified that the condition is also necessary.

The equation (19) is a relation between the various canonical variables, of which any $2n$ can be taken as independent. There will be four sets of these $2n$ independent variables:

$(q_1, \ldots, q_n;\ p_1, \ldots, p_n)$; $(\mathrm{Q}_1, \ldots, \mathrm{Q}_n;\ \mathrm{P}_1, \ldots, \mathrm{P}_n)$;
$(q_1, \ldots, q_n;\ \mathrm{P}_1, \ldots, \mathrm{P}_n)$; $(\mathrm{Q}_1, \ldots, \mathrm{Q}_n;\ p_1, \ldots, p_n)$.

The most important case is the selection of the variables

$$(q_1, \ldots, q_n;\ \mathrm{P}_1, \ldots, \mathrm{P}_n). \tag{21}$$

Then we must consider the equation (19) as an identity in q_r, P_r, so that we must equate separately the coefficients of each $\dot{q}_r$ and $\dot{P}_r$, as well as the term independent of these.

Now (19) can be written as:

$$\sum_r p_r\dot{q}_r - H(q_r, P_r, t) = \sum_r P_r\dot{Q}_r - K(q_r, P_r, t) + \frac{df}{dt}$$

$$= \frac{d}{dt}\sum_r P_rQ_r - \sum_r Q_r\dot{P}_r - K(q_r, P_r, t) + \frac{df}{dt}$$

$$= -\sum_r Q_r\dot{P}_r - K(q_r, P_r, t) + \frac{dF}{dt}, \tag{22}$$

where $F = f + \sum_r P_r Q_r$ is also an arbitrary function of the canonical variables q_r, P_r, and of t. From (22) we get then:

$$p_r = \frac{\partial F}{\partial q_r}, \quad Q_r = \frac{\partial F}{\partial P_r}, \quad (r = 1, 2, \cdots, n),$$

$$H(q_r, P_r, t) = K(q_r, P_r, t) - \frac{\partial F}{\partial t}. \tag{23}$$

In the particular case when H does not contain t explicitly, we can, in general, choose F so that it also does not contain t explicitly. The transformation equations (23) therefore become:

$$P_r = \frac{\partial F}{\partial q_r}(q_r, P_r), \quad Q_r = \frac{\partial F}{\partial P_r}(q_r, P_r), \quad (r = 1, \cdots, n);$$

$$H(q_r, P_r) = K(q_r, P_r). \tag{24}$$

We see that in this particular case, the Hamiltonian is invariant, which is evident since it represents the total energy.

2·6. *Integration of the Canonical Equations.*

In this paragraph we write ϕ_r for Q_r and α_r for P_r in order to conform to the usual notation in quantum theory. As we have seen, in order to integrate the canonical equations, we must choose a transformation such that the transformed Hamiltonian does not depend on $q_1, \cdots, q_n$ and t, but depends only on $\alpha_1, \cdots, \alpha_n$. The transformation equations (23) § 2·5 are now

$$p_r = \frac{\partial F}{\partial q_r}(q_r, \alpha_r, t), \quad \phi_r = \frac{\partial F}{\partial \alpha_r}(q_r, \alpha_r, t), \tag{1}$$

$$H(q_r, \alpha_r, t) = K(\alpha_r) - \frac{\partial F}{\partial t}(q_r, \alpha_r, t).$$

The new canonical equations of Hamilton would become:

$$\dot{\phi}_r = \frac{\partial K}{\partial \alpha_r}, \quad \dot{\alpha}_r = -\frac{\partial K}{\partial \phi_r} = 0, \tag{2}$$

so that all the α's are constant:

$$\alpha_1 = c_1, \quad \alpha_2 = c_2, \quad \cdots, \alpha_n = c_n; \tag{3}$$

and the new Hamiltonian K, being a function of the α's is also constant:

$$\mathrm{K}\,(\alpha_1, \cdots, \alpha_n) = \text{Constant} = \mathrm{E}. \tag{4}$$

Thus $\frac{\partial \mathrm{K}}{\partial \alpha_r}$ is also constant, and we find

$$\dot{\phi}_r = \frac{\partial \mathrm{K}}{\partial \alpha_r} = \omega_r \text{ (say)}, \tag{5}$$

i.e.,

$$\phi_r = \omega_r t + \beta_r, \qquad (r = 1, 2, \cdots, n). \tag{6}$$

Expressions (3) and (6) are the integrals of the canonical equations of motion, and they contain the $2n$ arbitrary constants c_r and β_r.

It remains for us now to determine the transformation function F. We set

$$\mathrm{F}\,(q_r, \alpha_r, t) = \mathrm{E}t + \mathrm{U}\,(q_r, \alpha_r, t), \tag{7}$$

and get on account of E depending on α_r only,

$$p_r = \frac{\partial \mathrm{F}}{\partial q_r} = \frac{\partial \mathrm{U}}{\partial q_r}, \; \phi_r = \frac{\partial \mathrm{F}}{\partial \alpha_r} = \frac{\partial \mathrm{E}}{\partial \alpha_r} t + \frac{\partial \mathrm{U}}{\partial \alpha_r} = \omega_r t + \frac{\partial \mathrm{U}}{\partial \alpha_r} \tag{8}$$

From (1), (4) and (8), we obtain therefore

$$\mathrm{H}\left(q_r, \frac{\partial \mathrm{U}}{\partial q_r}, t\right) + \frac{\partial \mathrm{U}}{\partial t} = 0, \tag{9}$$

which is Jacobi's equation.

We have consequently the theorem given by Jacobi: " If the complete integral U $(q_1, \cdots, q_n, \alpha_1, \cdots, \alpha_n, t)$ of the equation (9) has been found which contains the n arbitrary independent constants $\alpha_1, \cdots, \alpha_n$, then Hamilton's canonical equations of motion have the $2n$ integrals:

$$\frac{\partial \mathrm{U}}{\partial \alpha_r} = \phi_r - \omega_r t = \beta_r, \; \frac{\partial \mathrm{U}}{\partial q_r} = p_r, \quad (r = 1, \cdots, n), \tag{10}$$

where $\beta_1, \cdots, \beta_n$ are new arbitrary constants. "

In the particular case when H does not contain t explicitly, the Jacobi equation (9) can be simplified by writing, as in (6) § 2.5:

$$\mathrm{U} = \mathrm{S}\,(q_r) - \mathrm{E}t, \tag{11}$$

so that

$$p_r = \frac{\partial \mathrm{U}}{\partial q_r} = \frac{\partial \mathrm{S}}{\partial q_r}, \frac{\partial \mathrm{U}}{\partial t} = -\mathrm{E}. \tag{12}$$

Thus (9) becomes

$$H\left(q_1, \cdots, q_n, \frac{\partial S}{\partial q_1}, \cdots, \frac{\partial S}{\partial q_n}\right) = E. \tag{13}$$

A complete integral of this equation containing n arbitrary constants $a_1, \cdots, a_n$ will be $S\,(q_1, \cdots, q_n, E, a_1, \cdots, a_n)$, so that the function U in (11) will contain $n + 1$ constants $a_1, \cdots a_n$, E. We write therefore

$$E = a_1, \tag{14}$$

Then

$$U = S\,(q_1, \cdots, q_n, E, a_2, \cdots, a_n) - Et,\ F = Et + U = S\,(q, a), \tag{15}$$

so that from (10)

$$\frac{\partial U}{\partial E} = -t + \frac{\partial S}{\partial E} = \beta_1, \frac{\partial U}{\partial a_2} = \frac{\partial S}{\partial a_2} = \beta_2\,, \cdots, \frac{\partial S}{\partial a_n} = \beta_n. \tag{16}$$

The $(n-1)$ equations $\frac{\partial S}{\partial a_2} = \beta_2, \cdots, \frac{\partial S}{\partial a_n} = \beta_n$ give the form of the orbit, and the equation $\frac{\partial S}{\partial E} = \frac{\partial S}{\partial a_1} = \beta_1 + t$ gives the time at which the system has a certain position in the orbit.

We have also

$$\phi_r = \frac{\partial F}{\partial a_r} = \frac{\partial S}{\partial a_r}, \quad (r = 1, 2, \cdots, n). \tag{17}$$

If orthogonal co-ordinates are used, the function $H\,(p_r, q_r)$ contains the p_r as squares only, so that the Hamilton-Jacobi equation $H\left(\frac{\partial S}{\partial q_r}, q_r\right) = a_1$ contains the $\frac{\partial S}{\partial q_r}$ only as squares.

In the applications with which we shall be concerned in this book, it will be found that it is possible to separate the Hamilton-Jacobi equation into parts each of the form

$$\begin{gathered}\left(\frac{\partial S}{\partial q_1}\right)^2 = f_1\,(q_1, a_1, \cdots, a_n)\\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ \left(\frac{\partial S}{\partial q_n}\right)^2 = f_n\,(q_n, a_1, \cdots, a_n),\end{gathered} \tag{18}$$

such that if each part is satisfied, the equation (10) is satisfied as a whole. Any one equation in (18) contains only one of the co-ordinates q_r. If this separation can be effected, the complete integral of the Hamilton-Jacobi equation is then given by integrating the total differential equation

$$dS = \frac{\partial S}{\partial q_1} dq_1 + \frac{\partial S}{\partial q_2} dq_2 + \cdots + \frac{\partial S}{\partial q_n} dq_n$$

$$= \sqrt{f_1(q_1, a_r)}\, dq_1 + \sqrt{f_2(q_2, a_r)}\, dq_2 + \cdots + \sqrt{f_n(q_n, a_r)}\, dq_n,$$

which gives on integrating

$$S = \int \{f_1(q_1, a_r)\}^{\frac{1}{2}} dq_1 + \cdots + \int \{f_n(q_n, a_r)\}^{\frac{1}{2}} dq_n, \tag{19}$$

where a_r stands for all the $a_1, a_2, \cdots a_n$. The orbit and the time can then be determined from the equations (13).

2·7. *Periodic Motions and the Action and Angle Variables.*

In the present volume we shall be concerned only with periodic systems. Now, this periodicity is of two kinds. Either q oscillates to and fro periodically in time between two fixed limits, or the corresponding p is a periodic function of q. The first case is spoken of as one of liberation and is illustrated by the oscillating (simple) pendulum. The second case is that of rotation and is typified by the conical pendulum.

For convenience, we confine ourselves at first to systems with one degree of freedom. Then we choose the new variable ϕ of the last section in such a way that it increases by 1 during one period of the motion. We denote it then by w and the corresponding conjugate variable by J. w is then called an "angle variable" and J an "action variable". As w and J are particular examples of the variables ϕ_r, a_r used in the last section, we see that the transformation function S will depend on q, J, so that $S = S(q, J)$.

Then from equation (17) § 2·6, *viz.*, $\phi = \frac{\partial S}{\partial a_r}$, we get on writing w for ϕ_r and J for a_r,

$$w = \frac{\partial S}{\partial q}(q, J), \tag{1}$$

and

$$p = \frac{\partial S}{\partial q}(q, J). \tag{2}$$

The differential coefficient of w along the path is then

$$\frac{\partial w}{\partial q} = \frac{\partial}{\partial q}\left(\frac{\partial S}{\partial J}\right) = \frac{\partial}{\partial J}\left(\frac{\partial S}{\partial q}\right).$$

Then

$$\oint dw = \oint \frac{dw}{dq} dq = \frac{\partial}{\partial J} \oint \frac{\partial S}{\partial q} dq = \frac{\partial}{\partial J} \oint p\,dq, \tag{3}$$

where the symbol $\oint$ denotes that the integration is to be extended over one period, *i.e.*, in the case of liberation, over one back and forward motion of q, and in the case of a rotation, over a path of length 2π.

That the period of w shall be 1 therefore implies that $\oint dw = 1$, consequently

$$\frac{\partial}{\partial J}\oint \frac{\partial S}{\partial q}\,dq = \frac{\partial}{\partial J}\oint p\,dq = 1. \tag{4}$$

This can be satisfied if we take

$$\oint \frac{\partial S}{\partial q}\,dq = \oint p\,dq = J, \tag{5}$$

or, in other words, J is equal to the increase of S during one period.

The above analysis gives the meaning and properties of the action and angle variables. The question in an actual case is how to determine w, J. For a system with one degree of freedom we have the Hamiltorian H given as a function of some canonical variables p, q. The transformation function S (q, α) is then determined by integration of the Hamilton-Jacobi equation $H\left(\frac{\partial S}{\partial q}, q\right) = \alpha$. Then we find the value of

$$J = \oint \frac{\partial S}{\partial q}\,dq.$$

Obviously J is a function of α, and inversely, we can find α as a function of J. When we substitute this value of α in S (q, α), we get S as a function of q, J.

By means of the transformation

$$p = \frac{\partial S(q, J)}{\partial q},$$

$$w = \frac{\partial S(q, J)}{\partial J},$$

p and q will become periodic functions of w with the period 1, and H will be a function α of J alone, *i.e.*, $H = H(J)$. The canonical equations of motion are

$$\left.\begin{aligned} \frac{dJ}{dt} &= -\frac{\partial H}{\partial w} = 0, \quad \therefore J = \text{Constant}, \\ \frac{dw}{dt} &= \frac{\partial H}{\partial J} = \nu, \quad w = \nu t + \beta, \end{aligned}\right\} \tag{6}$$

where ν is a constant. Since we have chosen w so that it increases by 1 during each period of the motion, it follows that H is a function which increases continuously with J. Consequently ν must be a positive number; it is equal to the number of periods in unite time, or the frequency of the motion.

If the system has many degrees of freedom, say n, then we assume that each of the q_r has a periodic character of one of the two kinds explained above. We also assume, as frequently happens in quantum-theoretical applications, that the Hamilton-Jacobi equation may be solved by separation of the variables, *i.e.*, as in (19) § 2·6, that the transformation function S breaks up into a sum of n terms each of which depends on a single q_r; *i.e.*, in the case of cyclic co-ordinates:

$$S = S_1(q_1, a_1, \cdots, a_n) + \cdots + S_n(q_n, a_1, \cdots, a_n), \tag{7}$$

so that

$$p_r = \frac{\partial S_r}{\partial q_r}, \quad (r = 1, \cdots, n), \tag{8}$$

and p_r is a function of q_r alone. Therefore each integral

$$J_r = \oint p_r \, dq_r \quad (r = 1, \cdots, n) \tag{9}$$

taken over a period of q_r is constant, depending only on the constants $a_1, \cdots, a_n$, as is evident from (7) and (8). Each of the J_r is thus a function of $a_1, \cdots, a_n$, and inversely therefore each of the a_r is a function of $J_1, \cdots, J_n$. Particularly a_1, *i.e.*, H is a function of $J_1, \cdots, J_n$. Each of the S_r is then a function of q_r, $J_1, \cdots, J_n$.

We can now introduce the angle variables w_r conjugate to the action variables J_r by the equations

$$w_r = \frac{\partial S}{\partial J_r} = \sum_{k=1}^{n} \frac{\partial S_k}{\partial J_r}, \qquad (r = 1, \cdots, n). \tag{10}$$

It is then not difficult to show that the q_r's are multiply periodic functions of the w_r's, in the sense that *e.g.*, $q_1(w_1, \cdots, w_n)$ goes through one period when w_1 is increased by 1 while the other w's remain unaltered; the remaining q's $q_2, \cdots, q_n$ may also depend on w_1, but they return to the initial values without going through a complete period. Similarly the function $q_2(w_1, w_2, \cdots, w_n)$ goes through one period when w_2 is increased by 1 while all the other w's remain unaltered; $q_1(w_1, w_2, \cdots, w_n)$, $q_3(w_1, w_2, \cdots, w_n)$, $\cdots$, $q_n(w_1, w_2, \cdots, w_n)$ return to their initial values without going through a complete period.

For a non-degenerate system, *i.e.*, a dynamical system in which the number of independent frequencies $\nu_1, \cdots, \nu_n$ is equal to the number of degrees of freedom, the action variables $J_1, \cdots, J_n$

are invariant. This theorem was proved by J. M. Burgers in 1917. The action variables $J_1, \cdots, J_n$ are therefore also called the " adiabatic invariants ".

If the system is degenerate, *i.e.*, if the number of independent frequencies is less than n, it is possible, by a linear transformation, to find new action variables J_r' and angle variables w_r' equal in number to the independent frequencies. It is then only these new J_r' which are adiabatic invariants.*

---

* For a complete proof of these theorems, as well as a detailed account of the subject-matter of this section, the reader may consult Max Born's *Mechanics of the Atom,* Chapter II, or G. Birtwistle's *Quantum Theory of the Atom,* Chapter VII.

CHAPTER III.

STRUCTURE OF THE ATOM.

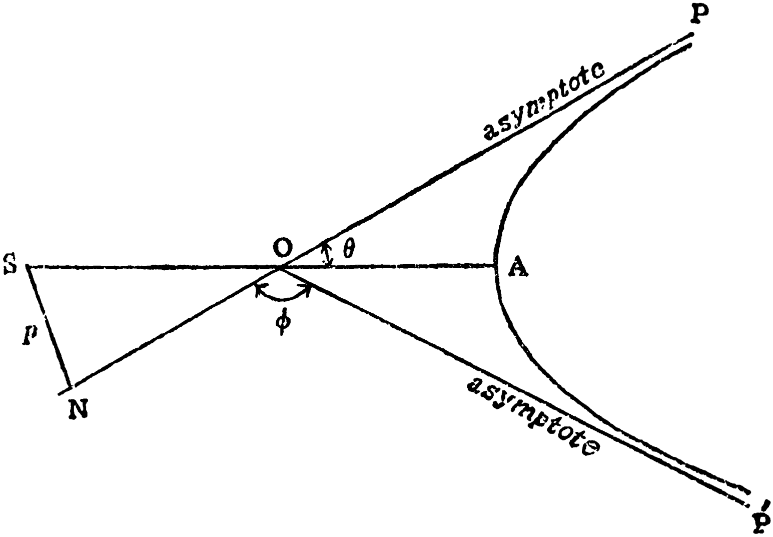
3·1. *Rutherford's Investigations.*

SINCE the eighteen-nineties, when J. J. Thomson performed his experiments on the discharge of electricity through gases, it has been known that the atom is not a simple indivisible unit, assumed previously, but is a complex structure consisting of still smaller particles. One kind of these particles were discovered by Thomson, and are now called " electrons ". The rest mass m of an electron is estimated to be about $9 \cdot 04 \times 10^{-28}$ grams, and it carries a negative charge of electricity $-e$, where $e = 4 \cdot 77 \times 10^{-10}$ E.S.U. An electric charge less than e has never been observed, and all experimental evidence leads to the conviction that it is the smallest amount of electric charge that exists in nature, all other charges being its integral multiples.

In the early years of the present century, Rutherford and his collaborators performed experiments to determine the nature of atomic structure. These experiments consisted in their main outline in letting a dense current of swift α-particles pass through a thin metal plate, and observing the deflections in their directions of motion. It was found that a small fraction of the α-particles were able to be deflected through a large angle, more than 90°. In the year 1911, Rutherford pointed out that these large deflections would be due to the α-particle passing through a strong electric field existing within the atom. Thus, Rutherford was led to the conclusion that the atom consists of a positively charged " nucleus " of very small dimensions in which practically all the mass of the atom is concentrated. As the least value of the charge is e, the positive charge on the nucleus can only be $+Ne$, where N is an integer. However, since a normal atom is electrically neutral, Rutherford assumed at the same time that the nucleus is surrounded by the number N of electrons, arranged about it at various distances.

The large deflections of some of the α-particles can, on this nuclear hypothesis, be accounted for by supposing that they pass through the atoms in their path and enter the intense electric field in the neighbourhood of the nucleus, thus suffering deflection from their rectilinear path. It is obvious that this deflection could not be caused by the surrounding electrons, since the mass of an α-particle is more than 7,000 times that of an electron, and its velocity is immense, being about one-tenth of the velocity of light.

The value of the integer N in the charge N*e* on the nucleus can be deduced readily from the observed value of the deflection angle, if it is assumed that the forces between the α-particle and the nucleus are repulsive, and follow the law of inverse squares. It is known from elementary dynamics that in this case the α-particle describes a hyperbola round the nucleus as the external focus.



Suppose that the nucleus of the atom is S, possessing a charge + N*e*, and that the mass of the α-particle is M and its charge is + E. As a matter of fact, an α-particle is a helium atom doubly ionised, and we know that its mass is about four times that of the hydrogen atom and its charge is + 2*e*. Let the α-particle enter the atom in the direction PO such that it just misses the nucleus S by a distance *p*. This means that the distance SN from the nucleus S on the tangent at P is *p*. Let the velocity of the particle at P be *v*, and suppose that the particle describes part of the hyperbola PAP′, so that on escaping the atom its direction of motion is OP′. OP and OP′ are thus the two asymptotes and O the centre of the hyperbola. The deflection in the direction of the particle is ϕ = angle NOP′. If the lengths of the semi-transverse and semi-conjugate axes are a and b respectively, and if 2θ is the angle between the two asymptotes, then we know that $\tan\theta = \frac{b}{a}$. But $\phi = \pi - 2\theta$, so that $\theta = \frac{1}{2}(\pi - \phi)$, therefore

$$\tan\left(\frac{\pi}{2} - \frac{\phi}{2}\right) = \frac{b}{a},$$

i.e.,

$$\cot\frac{\phi}{2} = \frac{b}{a}. \tag{1}$$

But, if ϵ is the eccentricity given by $\epsilon^2 - 1 = \frac{b^2}{a^2}$,

$$\begin{aligned} p &= \mathrm{OS}\sin\theta = a\epsilon\sin\theta \\ &= a\epsilon\,\frac{b}{\sqrt{a^2+b^2}} = \sqrt{a^2+b^2}\cdot\frac{b}{\sqrt{a^2+b^2}} = b. \end{aligned} \tag{2}$$

Moreover, if the distance SP is r_1, we have from the dynamical theory

$$v^2 = \mu\left(\frac{1}{a} - \frac{2}{r_1}\right), \tag{3}$$

where μ is the force on the particle per unit mass at a unit distance. But compared to the dimensions of both the atomic nucleus S and the α-particle, the distance r_1 is so large that we can neglect the term $\frac{2}{r_1}$. Then

$$v^2 = \frac{\mu}{a}. \tag{4}$$

Now the force between the nucleus and the α-particle at any point of the orbit is $\frac{\mathrm{N}e\mathrm{E}}{r^2}$, so that the force on the particle per unit mass is

$$\mathrm{F} = \frac{\mathrm{N}e\mathrm{E}}{\mathrm{M}r^2}. \tag{5}$$

Thus, on account of $\mathrm{F} = \frac{\mu}{r^2}$, we obtain:

$$\mu = \frac{\mathrm{N}e\mathrm{E}}{\mathrm{M}}. \tag{6}$$

From (4) and (6) we get therefore

$$v^2 = \frac{\mathrm{N}e\mathrm{E}}{\mathrm{M}a},\ i.e.,\ a = \frac{\mathrm{N}e\mathrm{E}}{\mathrm{M}v^2}. \tag{7}$$

Substituting (2) and (7) in (1), we get

$$\cot\frac{\phi}{2} = \frac{b}{a} = \frac{pMv^2}{NeE},$$

or

$$N = \frac{pMv^2}{eE}\tan\frac{\phi}{2}. \tag{8}$$

It was at first deduced from observations of the deflection angle ϕ that the value of N was equal to about half the atomic weight. In 1913 Van den Brock suggested that the value of N was exactly equal to the atomic number, *i.e.*, to the serial number of the element when all the elements are arranged in the periodic table of Mendeleef in order of the increasing atomic weight. This view was advocated by Bohr who adopted it for his quantum theory and it is now universally accepted.

The fundamental result of Rutherford's experiments is, therefore, that each atom consists of two parts. The centre of the atom is occupied by a "nucleus" which is much smaller than the whole atom, and which carries practically the whole weight of the atom. The nucleus, moreover, is positively charged, the magnitude of the charge depending on the atomic number N of the element, being in fact equal to Ne. The second part of the atom, the extra-nuclear part, consists of a number of peripheral electrons, their number being equal to the atomic number N.

As the dimensions of the nucleus are exceedingly smaller than those of the atom, the internal structure of the nucleus will not affect sensibly the distribution of the peripheral electrons. Consequently, as pointed out by Bohr, the ordinary chemical and physical properties of the elements will depend entirely on the number and distribution of the extra-nuclear electrons.

The internal structure of the nucleus is responsible for the phenomena of Isotopy and Radioactivity.

3·2. *Rutherford's Atom-Model.*

At the time of Rutherford's investigations, and even as late as 1932, only two kinds of fundamental particles were known; (1) the electron, carrying a mass m and a negative charge $(-e)$, (2) and the ionised hydrogen atom, or the hydrogen-nucleus, with a mass, 1847 m approximately, and carrying a positive charge $(+e)$. This hydrogen-nucleus is now given the name "proton". The proton was until few years ago the smallest known positively charged particle. It was therefore considered that all matter must be constructed of the two fundamental particles, electrons and protons.

Rutherford applied this conception to his nuclear model, and gave the following account of the structure of the atoms of all the chemical elements in the Periodic Table.

Let an atom belong to the chemical element of atomic number N, and atomic weight w, where it is known that except for hydrogen, *i.e.*, $N = 1$, w is always greater than N. From his investigations Rutherford knew that the nucleus has a charge $+Ne$. Moreover, the weight of the atom is w times the weight of the hydrogen atom, and this whole weight is concentrated in the nucleus. Since the proton is much more heavier than the electron, we can neglect the weight of the electrons when considering atomic weights. The weight of the hydrogen atom is, therefore, taken to be the weight of its nucleus, *i.e.*, that of a proton. The nucleus of the heavier atom must therefore contain a number w of protons, to give the atom the necessary weight w. This would, however, give the nucleus a charge $(+we)$ greater than the actual charge $(+Ne)$. This charge can be reduced only when the nucleus contains a number $(w - N)$ of electrons.

Thus, according to Rutherford, the given atom has a nucleus consisting of w protons and $(w - N)$ electrons. Around this nucleus are distributed a number N of peripheral electrons.

For example, the atom of hydrogen has $N = 1$, $w = 1$. The hydrogen-nucleus consists, therefore, of one proton only, and the atom has one peripheral electron revolving round the proton.

For helium, $N = 2$, $w = 4$. The helium nucleus consists therefore of 4 protons and $4 - 2 = 2$ electrons. This is the well-known alpha particle. Around this nucleus there are two peripheral electrons.

Similarly for all the atoms of the Periodic Table.

In this account of the structure of atoms we have taken for granted that the atomic weight w is an exact integer. But this is not always the case. The weight of the heavier atoms is, in general, not an exact integral multiple of the weight of the hydrogen atom. For instance, the atomic weight of the neon atom is 20·17, and that of the chlorine atom is 35·44. To account for this discrepancy, it was assumed that there exist variants of the same elements, *i.e.*, atoms, with the same nuclear charge and the same atomic number, but with different atomic weights. These variants of the same element are called " Isotopes ". The actual existence of the isotopes was proved by the experiments of F. W. Aston, who designed his " Mass-Spectrograph " to detect them. Two isotopes of the same element would have the same spectral and chemical peoperties, since these properties would depend on the number of the peripheral electrons, or what is the same thing, on the nuclear charge only. And this nuclear charge is the same for the two isotopes. They differ simply in the structure of their nuclei. Thus, chlorine has two Isotopes of atomic weights 35 and 37—exact integral multiples. The nucleus of the first isotope consists of 35 protons and 18 electrons, and there are 17 electrons round this nucleus. The nucleus of the second isotope consists of 37 protons and 20 electrons, and there are 17 electrons revolving round this nucleus. Chlorine, as found ordinarily, is a mixture of these two isotopes, and the atomic weight determined by chemical methods, is the mean atomic weight of the two isotopes.

Various isotopes of almost all the chemical elements, including hydrogen and oxygen, have been discovered. Some elements have more than two isotopes, thus tin has 11 and xenon 9 isotopes.

This picture of the nuclear structure suggested by Rutherford, prevailed practically unaltered till 1932. In that year, however, two more fundamental particles, *viz.*, the neutron and the positron, were discovered, placing the question of nuclear structure on entirely different foundations.

3·3. *The Neutron.*

The existence of neutrons was postulated by Dirac and by Pauli in 1931 on theoretical grounds, but was experimentally demonstrated by J. Chadwick in 1932. When swiftly moving α-particles, emitted from the radioactive polonium, strike a beryllium plate, it was found by Bothe and Becker in 1931 that a radiation is ejected which is of such enormous penetrating power that it can pass through thick metal plates. It could not therefore be the ordinary radiation. It was found further that this " beryllium radiation " consisted neither of a shower of protons, nor of α- and β-particles. Chadwick pointed out that this radiation could not even be the electromagnetic γ-rays, but that it must consist of uncharged particles moving with large velocities. He called them "neutrons," and determined experimentally that the mass of a neutron is very nearly equal to the mass of a proton. The collision of a beryllium atom with an α-particle producing a neutron can be represented by the equation :

$$Be_4^9 + He_2^4 \rightarrow C_6^{12} + n_0^1,$$

where the upper indices denote atomic weights, and the lower indices the atomic numbers (nuclear charges).

It was at first believed that the neutron is not an elementary particle, but composed of an electron and a proton. But this view has been given up, and the four particles, *viz.*, electron, proton, neutron and positron, are all considered to be elementary. The dimensions of the neutron are much smaller compared to the distance between the atomic nucleus and the peripheral electrons. It is evident therefore, that the neutron can easily pass through the atoms in the metal plate, the more so as it has no charge and is consequently unaffected by the electric fields of the nucleus and the electrons.

3·4. *The Positron.*

Dirac's relativistic-quantum mechanics published in 1928, made it theoretically evident that there exists the true counterpart of the electron, *i.e.*, a particle of mass equal to that of the electron but with a positive charge *e* (*cf.* Chapter XI, especially § 11·6). The actual experimental demonstration of the existence of such a particle was, however, long delayed, and it was only at the end of 1932 that Anderson and Blacket could discover it independently. These particles are called positive-electrons, or more shortly " positrons ".

The positrons are created when the highly penetrating cosmic rays fall on matter, or when fast moving particles strike an atomic nucleus. Their presence is detected by means of the Wilson chamber from the fact that their tracks are curved in the opposite direction to the tracks of the electrons. They have, however, an extremely large tendency of combining with the electrons, so that both the positron and the electron disappear and radiation is emitted. As remarked in the last section, the positron is an elementary particle, and one of the primary constituents of matter.

3·5. *Structure of the Nucleus.*

In Section 3·2 we saw that Rutherford considered an atomic nucleus to consist of protons and electrons, because these were the only two kinds of ultimate particles known at that time. Since then, however, as described in the last two sections, two more ultimate particles, *viz.*, the neutrons and the positrons, have been discovered. It was natural therefore, that physicists should consider a revision in our ideas of the nuclear structure.

This revision became inevitable when investigations on the hyperfine structure of spectral lines and the stability of radioactive elements, showed clearly that free electrons cannot be a constituent of the nucleus.

In 1932, Heisenberg suggested that atomic nucleii are composed of neutrons and protons only, there being no free electrons. The hydrogen-nucleus consists of one proton. The helium nucleus (α-particle) consists of two neutrons and two protons. $He_2^4 \rightarrow 2\, n_0^1 + 2\, p_1^1$. This is a particularly stable configuration, and therefore α-particles enter as fundamental particles in the combination of heavier nucleii. Thus the nucleus of chlorine, of atomic weight 35 and atomic number 17, would be composed of 8 α-particles, one proton and two neutrons:

$$Cl_{17}^{35} \rightarrow 8\, He_2^4 + 2\, n_0^1 + p_1^1,$$

where we assume that the proton has " atomic number " 1.

Now, if a neutron is added to the nucleus of the atom of any element, it does not increase the nuclear charge, but only the weight of the atom, thus giving rise to an isotope of that particular element. If more neutrons are added, further isotopes of the same element are produced. The number of isotopes is determined by certain conditions of stability. The neutron thus explains the formation of the isotopes as well.

This account of the nature of matter is by no means complete. It is only one side of the picture, as it were. In Chapter VII we shall see that matter has not only a corpuscular aspect but a wave-aspect also. This dualism is a universal phenomenon in the domain of modern physics inasmuch as it is an integral and experimentally established fact in the theory of light as well. We shall devote the next chapter to an exposition of this dual nature of light.

CHAPTER IV.

THE QUANTUM THEORY OF RADIATION.

4·1. *The Dual Nature of Light, and Einstein's Light-quantum.*

THE optical phenomena generally known in the time of Newton and Huygens, were those of rectilinear propagation and of reflection and refraction. It is well known that Newton proposed a corpuscular theory whereas Huygens put forward a wave-theory, to account for these phenomena. The controversy between the followers of Newton and those of Huygens was, however, mainly a philosophical one. The facts themselves, so far as they were known in those days, could be explained equally well on both the corpuscular and the wave hypotheses.

Experimental investigations during the early part of the 19th century revealed the hitherto undetected phenomena of interference and diffraction. It was found that these phenomena could be explained only on the wave-theory. The particle picture failed for them completely, and fell into disrepute. When Maxwell published his electromagnetic theory, and when this last theory was confirmed later by Hertz's discovery of the electromagnetic-waves, the corpuscular hypothesis was completely obliterated, and the wave-theory reigned supreme in the domain of optics.

By a strange irony, Hertz himself was the discoverer of yet another phenomenon which was to challenge the absolute supremacy of the wave-theory, and which ultimately brought the corpuscular view again into the picture. There is this difference, however, that whereas in the 18th century, the two views were rival ones in the sense that one of them could be chosen and the other discarded at will, they are complementary now in being both essential to a description of reality.

This latter discovery of Hertz—called the " Photo-electric Effect "—defied all attempts at an explanation on the electromagnetic or wave-theory. Einstein accounted for it in 1905 by going back to the corpuscular hypothesis. This conception of Einstein's is, however, not exactly the same as the old one of Newton.

Einstein took his cue from an earlier idea of Planck's—the idea of the quantum of action put forward in 1900 (see Section 4·3). He assumed that the energy of the incident radiation does not get scattered in waves in all directions, but that it remains concentrated in a stream of moving particles or "light quanta". Each quantum of light—or "photon", as it is now generally called—moves with the velocity c of light. We shall prove in § 4·2 that the energy of each photon is a constant multiple of the frequency.

Oddly enough, Einstein's theory of light-quanta did not receive much attention for a number of years, until in 1923 it was found necessary for the explanation of yet another discovery—the Compton Effect (§ 4·5). Since then it has made great strides, and has as firmly established itself now as the wave-theory.

When Einstein introduced it in 1905, the conception of the light-quantum was a mere hypothesis which explained the photo-electric phenomena. The experiments of Bothe and Geiger and those of Compton and Simon in 1925, provided more or less direct evidence for the existence of the photons.

All this does not mean, however, that the corpuscular theory of light does away with the wave-theory. We have seen that the phenomena of interference and diffraction require that light must behave as waves. We must retain therefore both the corpuscular and the wave conceptions. They are both necessary and they supplement each other.

This duality is one of the most fundamental facts on which the structure of Quantum Mechanics is built. In § 7·1 we shall see that there is an analogous duality in the nature of matter. Duality is therefore a characteristic of our universe. In Chapter X, we shall see how modern quantum mechanics attempts at reconciling these two contradictory conceptions of corpuscles and waves.

4·2. *Relativity Theory of the Photon, and Derivation of Planck's Law* $\epsilon = h\nu$.

We shall use the experimental fact of the dual nature—the corpuscular and the wave-nature—of light in conjunction with the relativity theory to establish important relations between the corpuscular quantities, energy and momentum, and the wave-quantities, frequency and wave-length.

We have seen in § 1·7 that the momentum vector $\vec{p}$ and energy ϵ of a particle are given by

$$\vec{p} = m\,\vec{v}, \tag{1}$$

and

$$\epsilon = m\,c^2, \tag{2}$$

where m is the mass, v the velocity of the particle, and c the velocity of light. If m_0 is the rest mass, then

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{3}$$

This equation shows that v can never be greater than c. No material particle can ever have even the limiting value c, for otherwise its mass m would become infinite. There is, however, one case when even for $v = c$, the mass m can remain finite, and this case occurs only when the numerator m_0 also vanishes. Thus there can be only one kind of particle with the velocity c, and this has zero rest mass. This particle is Einstein's light-quantum or photon.

Now, since $v = c$ for a photon, we get from (1) and (2)

$$p = mv = \frac{\epsilon}{c^2}\cdot v = \frac{\epsilon}{c^2}\cdot c,$$

i.e.,

$$p = \frac{\epsilon}{c}. \tag{4}$$

Let λ be the wave-length and ν the frequency in the corresponding wave-picture, then the wave-number k is defined in optics as the reciprocal of the wave-length, *i.e.*, $k = \frac{1}{\lambda}\cdot$ But $\lambda = \frac{c}{\nu}$ so that we get

$$k = \frac{\nu}{c}. \tag{5}$$

Comparing the two equations (4) and (5), we find that the momentum p corresponds to the wave-number k, and the energy ϵ corresponds to the frequency ν. But $\vec{p}$ is a 3-dimensional vector having components p_x, p_y, p_z. To get the desired relation between the corpuscular and wave-pictures, we must therefore assume that the wave-number k is also a vector quantity having the components k_x, k_y, k_z, which determine the direction of propagation of the light-waves. This means that the direction of motion of the photons in the particle picture coincides with the direction of propagation of the light waves in the wave-picture.

The required relations between the corpuscular quantities p_x, p_y, p_z, ϵ and the wave-quantities k_x, k_y, k_z, ν are obtained by an application of the condition that they must be invariant for a Lorentz-transformation.

In the wave-picture suppose that the system of waves with wave-length λ and frequency ν is propagated in a direction making angles α, β, γ with the co-ordinate axes. The wave-equation is then known to be

$$\psi = \psi_0 \cos 2\pi \left(\frac{x \cos \alpha + y \cos \beta + z \cos \gamma}{\lambda} - \nu t \right), \tag{6}$$

where ψ_0 denotes the amplitude of the waves.

We have seen already that the wave-number $k = \frac{1}{\lambda}$ must be assumed a vector quantity, so that its components are given by

$$k_x = \frac{\cos \alpha}{\lambda}, \quad k_y = \frac{\cos \beta}{\lambda}, \quad k_z = \frac{\cos \gamma}{\lambda}. \tag{7}$$

Then equation (6) becomes

$$\psi = \psi_0 \cos 2\pi \, (xk_x + yk_y + zk_z - \nu t). \tag{8}$$

The phase of vibration ϕ is therefore

$$\phi = 2\pi \, (xk_x + yk_y + zk_z - \nu t). \tag{9}$$

This phase of vibration has evidently the dimensions $\left[\frac{\text{length}}{\text{length}}\right] = [\text{length}]^0$, *i.e.*, it has no dimensions, and is consequently an invariant quantity. From vector analysis we know that the scalar product of two vectors is also an invariant quantity, being independent of the choice of axes. The expression

$$xk_x + yk_y + zk_z - \nu t \tag{10}$$

must therefore be a scalar product of two 4-dimensional vectors, since it is invariant. But we know that the space-time vector has the components

$$x, y, z, ict, \tag{11}$$

so that the other four-vector of the product (10) must have components

$$k_x, k_y, k_z, \frac{i\nu}{c}. \tag{12}$$

Now the expression

$$x^2 + y^2 + z^2 - c^2 t^2, \tag{13}$$

is also a scalar product of the vector (11) with itself. The two scalar products (10) and (13) should therefore behave similarly under a Lorentz-transformation, and are therefore invariant together.

We see thus that the phase (10) remains invariant under a Lorentz-transformation only when it is characterised by the 4-dimensional vector (12).

On the other hand, we have seen in § 1·7 that both momentum and energy together form a 4-dimensional vector with the components (5) § 1·7 :

$$p_1 = p_x = mv_x,\ p_2 = p_y = mv_y,\ p_3 = p_x = mv_x,\ p_4 = i\frac{\epsilon}{c} = imc.$$

We have thus two 4-dimensional vectors with components

$$\begin{aligned} &p_x,\ p_y,\ p_z,\ \frac{i\epsilon}{c}, \\ &k_x,\ k_y,\ k_z,\ \frac{i\nu}{c}. \end{aligned} \tag{14}$$

So that if there exists an invariant relation between the corpuscular quantities p_x, p_y, p_z, ϵ, and the wave-theory quantities k_x, k_y, k_z, ν, it must be of the form :

$$\frac{p_x}{k_x} = \frac{p_y}{k_y} = \frac{p_z}{k_z} = \frac{\epsilon}{\nu} = h, \text{ (say)}. \tag{15}$$

Since this relation has to be invariant and independent of the choice of co-ordinate axes, h must be an absolute constant. It is called Planck's constant (§ 4·3).

From (15) we have therefore

$$p_x = hk_x,\ p_y = hk_y,\ p_z = hk_z, \tag{16}$$

and

$$\epsilon = h\nu. \tag{17}$$

Instead of the three equations (16) we can write the one equation

$$p = hk = \frac{h}{\lambda}, \tag{18}$$

or, remembering that $\lambda = \frac{c}{\nu}$, we can write

$$p = \frac{h\nu}{c}. \tag{19}$$

The two equations (17) and (19) are consistent with the two equations (1) and (2), or with the equation (4), since they give

$p = \frac{\epsilon}{c}$. They show that the photon has energy $h\nu$ and momentum $\frac{h\nu}{c}$. Thus the same ray of light which, when considered as a wave, has frequency ν and wave-length λ, has the energy $h\nu$ and momentum $\frac{h\nu}{c}$ when considered as a particle.

Equation (17), *viz.*, $\epsilon = h\nu$, gives Planck's famous relation for the quantum of energy. It is the most fundamental relation of the quantum theory, and has been in fact the starting point of the whole subject. We have derived it here from the relativity theory on the experimental evidence that light behaves sometimes as particles and sometimes as waves. Actually, however, the development was in the reverse order. We shall see in the next section that Planck proposed it—as an arbitrary hypothesis, of course—long before the corpuscular behaviour of light was recognised. We shall also see what a revolutionary conception it was, in as much as it admitted discontinuity in natural processes. Based on this energy hypothesis of Planck's, Einstein developed his quantum theory of light, and derived the momentum relation $p = \frac{h\nu}{c}$ by substituting (17) in (4).

4·3. *Black-body Radiation, and Planck's Quantum Hypothesis.*

The quantum theory originated in an attempt to account for the characteristic properties of the heat radiated from a hot body. Let a hot body be maintained throughout its substance at a constant temperature, and suppose that there is a cavity in the interior of such a body. Inside the cavity there will be a continuous stream of radiation in all directions.

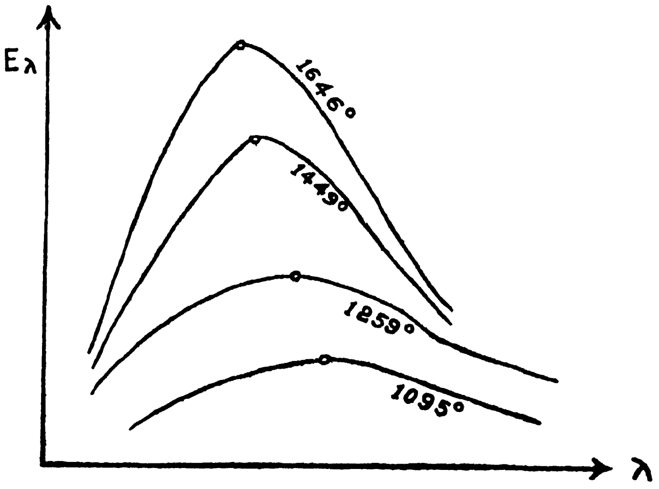
The problem is to determine the distribution of energy between the various wave-lengths of frequencies which make up the complete continuous spectrum of the black-body radiation.

In 1899, Lummer and Pringsheim solved this problem experimentally in the following way. A small hole is made in the hollow body raised to a uniform high temperature. The interior is seen to be incandescent, and it is well known that all bodies contained in such a furnace appear to be of the same colour. This colour is characteristic of the temperature, and is independent of the nature of the substance heated. The radiant energy consists of a vast range of mono-chromatic radiations, each present in a definite intensity, so that the total radiation has a definite spectral distribution. Lummer and Pringsheim measured the intensity of each radiation, and traced a curve representing the distribution of energy among the various wave-lengths. They found that there is a particular wave-length λ, intermediate between the upper and lower limits, where the energy of emission is a maximum.

Moreover, as expressed by Stefan's law

$$W = \sigma t^4, \tag{1}$$

where σ is Stefan's constant $= 5 \cdot 72 \times 10^{-5}$, and W is the total quantity of energy radiated per second in all directions, the total energy is increased as the temperature t is raised. At the same time, the wave-length corresponding to the maximum energy moves towards left, *i.e.*, it becomes shorter. These facts are described in the accompanying figure, where E_λ denotes the energy corresponding to the wave-length λ.



During the latter part of the nineteenth century, many attempts were made to account for these experimental facts on the basis of the classical theory. Kirchhoff had shown in 1860 that the amount of energy per unit volume, as well as the distribution of this energy among the various wave-lengths, is independent of the form and the size of the cavity and of the surrounding bodies, and depends only on the temperature. From the general principles of thermodynamics, Wien deduced the law that

$$E_\lambda = \frac{1}{\lambda^5} f(\lambda t), \tag{2}$$

where f is a function entirely independent of the nature of the substance, thus confirming Kirchhoff's results.

The principles of thermodynamics are, however, not sufficient for the determination of the analytical form of f, which can be obtained only when some special hypothesis is made as regards the mechanism of emission and absorption of energy by the vibrating electron.

Wien proposed that $f(\lambda t)$ should be taken as $a\,e^{-\frac{b}{\lambda t}}$, where a and b are constants, so that

$$E_\lambda = \frac{a}{\lambda^5} e^{-\frac{b}{\lambda t}}. \tag{3}$$

It was found, however, that this formula gave good results only for short wave-lengths, but failed for long wave-lengths.

At the same time, Rayleigh found that $f(\lambda t) = c\lambda t$, where c is a constant, and later Jeans proved that the classical electrodynamics could lead only to this form for the function f, so that

$$E_\lambda = \frac{c}{\lambda^5} \lambda t = \frac{ct}{\lambda^4}. \tag{4}$$

This formula again was found to be deficient in so far as it gave good results for long wave-lengths only, and failed for short wave-lengths.

Planck then discovered that the experimental facts are truly represented if $f(\lambda t)$ is taken as $c_1 \dfrac{1}{e^{\frac{c_2}{\lambda t}} - 1}$, so that the law of radiation is

$$E_\lambda = \frac{c_1}{\lambda^5} \frac{1}{e^{\frac{c_2}{\lambda t}} - 1}, \tag{5}$$

where c_1 and c_2 are certain constants.

If λ is small, $\frac{1}{\lambda}$ is large and therefore

$$\frac{1}{e^{\frac{c_2}{\lambda t}} - 1} \sim \frac{1}{e^{\frac{c_2}{\lambda t}}} = e^{-\frac{c_2}{\lambda t}},$$

so that for short wave-lengths Planck's formula (5) agrees with Wien's formula (3).

If λ is large, then $\frac{1}{\lambda}$ is small, so that neglecting higher powers we get

$$\frac{1}{e^{\frac{c_2}{\lambda t}}-1} \sim \frac{1}{\left(1+\frac{c_2}{\lambda t}\right)-1} = \frac{1}{\frac{c_2}{\lambda t}} = \frac{\lambda t}{c_2}.$$

Thus, for long wave-lengths Planck's formula (5) agrees with the Rayleigh-Jeans formula (4).

It was found, however, that the formula (5) could not be derived theoretically from the classical theory, so that Planck saw himself forced to break away from classical ideas. On 14th December 1900, he put forward the epoch-making hypothesis that the charged particle—usually called the oscillator, or vibrator,—which is the source of the monochromatic light, absorbs and emits energy only in discrete quanta. It changes its energy not continuously, as supposed in the classical theory, but by sudden jumps. The oscillator can have therefore only discrete amounts of energy E given by

$$E = n.h\nu, \quad (n = 1, 2, 3, \cdots), \tag{6}$$

where ν is the frequency of the radiation, and h is a universal constant, called Planck's constant. It is evident from (6) that h has the dimensions of energy divided by frequency, *i.e.*, of energy multiplied by time. It has therefore the character of an action. Its value has been determined to be

$$h = 6{\cdot}55 \times 10^{-27} \text{ erg} \times \text{sec.} \tag{7}$$

approximately. It is one of the most fundamental constants of physics and dominates the whole of atomic theory.

4·4. *Theory of the Photo-electric Effect.*

In 1887, Hertz let a beam of ultra-violet rays fall on an uncharged conductor, and found that under the influence of the rays the conductor became positively charged. After the discovery of the electrons it was recognised that the cause of this electrification was the ejection of electrons from the atoms of the conductor.

This effect is called the "photo-electric effect", and the electrons which are emitted are called the "photo-electrons". It is now generally known that photo-electrons are emitted from a large number of metals and from some gases, under the influence of X-rays, ultra-violet rays and even visible rays.

It is known, further that if the incident light is monochromatic, the number of electrons emitted per second is proportional to the intensity of the light. But the energy of the emitted electrons does not depend either on the intensity of the incident light or on the temperature of the metal. The only effect of the diminishing intensity is that fewer electrons are emitted. The energy, and consequently the velocity, of the photo-electrons is found to depend only on the frequency ν of the incident light. These velocities are found to range from zero, upto a certain maximum v which depends on the frequency ν. If the frequency ν of the incident light is less than a certain minimum value ν_0, no photo-electrons are emitted. The value of ν_0 is different for different metals. The velocity v of the photo-electron increases with the frequency ν of the incident light.

As already remarked, the classical wave-theory of light was quite unable to account for these facts of the photo-electric effect. In 1905 Einstein showed that they could be explained only when recourse is had to a corpuscular theory of light. He assumed that light of a definite frequency ν consists of particles having the energy $\epsilon = h\nu$, and momentum $p = \frac{h\nu}{c}\cdot$ This quantum of energy does not get scattered in spherical waves in all directions, but remains concentrated in a small region of space. In every elementary process of emission and absorption of light, energy can be transformed only by an amount $h\nu$.

On this hypothesis of the light-quanta, Einstein could easily account for the photo-electric phenomena. He pointed out that the emission of an electron from an atom is due to the absorption by the atom of one of these quanta. If ϵ is the energy required to remove the electron from the atom, then $h\nu - \epsilon$ is used in giving it the kinetic energy of motion after the emission, so that if E is the kinetic energy with which a photo-electron is ejected, we must have

$$E = h\nu - \epsilon. \tag{1}$$

If m is the mass and v the velocity of the emitted electron, then we have

$$\tfrac{1}{2}mv^2 = E = h\nu - \epsilon. \tag{2}$$

This equation gives us the velocity of the photo-electron, which is expelled under the influence of the light of a given wave-length. ϵ is constant for a given atom.

Since $\frac{1}{2} mv^2$ cannot be negative, we see from (2) that no photo-electron can be emitted unless $h\nu \geqslant \epsilon$, *i.e.*, unless

$$\nu \geqslant \frac{\epsilon}{h}, \quad \therefore \; \nu \geqslant \nu_0 \text{ where } \nu_0 = \frac{\epsilon}{h}. \tag{3}$$

ν_0 is the lowest frequency under which an electron can be just liberated, and then its velocity would be zero. ν_0 is therefore called the " threshold frequency ".

Combining the equations (2) and (3), we get :

$$\tfrac{1}{2}mv^2 = h\,(\nu - \nu_0). \tag{4}$$

The maximum value of the velocity v would be when $\nu_0 = 0$, *i.e.*, on account of (3) when $\epsilon = 0$, *i.e.*, when no work has to be done in liberating the electron, this maximum value being

$$v_{max.} = \sqrt{\frac{2h\nu}{m}}. \tag{5}$$

Thus the velocities of the emitted photo-electrons range from zero upto the maximum value (5).

If V is the potential difference, necessary to obtain electrons of energy $\frac{1}{2}\, m\, v^2$, then we can write

$$\tfrac{1}{2}mv^2 = eV, \tag{6}$$

where e is the electronic charge. From (6) and (2) :

$$\tfrac{1}{2}mv^2 = h\nu - \epsilon = eV. \tag{7}$$

This is Einstein's equation, and it was verified experimentally by R. A. Millikan in 1916.

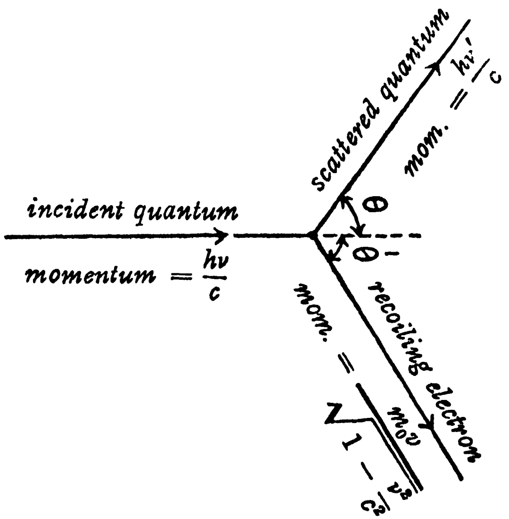
Einstein's theory of light-quanta thus explains all the features of the photo-electric effect in detail. The basic conception of this theory is Planck's law, that the interchanges of energy between atom and radiation can only occur discontinuously by quanta, and not continuously as assumed in the classical theory.

4·5. *Theory of the Compton Effect.*

In the year 1923, A. H. Compton discovered that when a beam of X-rays fell on a body, the frequency of the scattered radiation remained unaltered in the direction of the incident rays, but became lower in other directions, the frequency being in fact a function of the angle of deviation.

This change in the frequency, or the wave-length, of the scattered radiation is contradictory to the electro-dynamical theory of scattering given by J. J. Thomson. According to Thomson's theory the frequency of the scattered radiation must be the same as that of the original radiation. The classical theory therefore fails to account for the Compton effect. Consequently, Compton saw himself forced to adopt the light-quantum hypothesis put forward by Einstein several years ago, and found that the quantum theory gave results in agreement with experiments.

According to the classical theory each X-ray affects every electron in the matter traversed, and the scattering observed is due to the combined effect of all the electrons. Assuming the light-quantum hypothesis, Compton pointed out that a quantum of X-rays is not scattered by all the electrons in the radiator, but gives up all of its energy and momentum to some particular electron. This electron, in its turn, scatters the rays in a definite direction making the angle θ with the incident beam.



This deviation in the path of the quantum of radiation results in a change in its momentum. The rest of the momentum is spent in giving the electron a motion of recoil in some direction θ'. The principles of conservation of energy and momentum give the required values of the frequency ν' of the scattered radiation, and the recoil velocity v of the scattering electron.

This takes it for granted that these conservation principles are valid for the elementary quantised processes. Until a year or two ago this was in fact the general conviction among physicists. The experiments of Böthe and Geiger, and of Compton and Simon, were supposed to have established the universal validity of these principles.

But recently some observations have been made in connection with the radio-active disintegration of substances, which appear not to agree with the principle of conservation of energy. This is particularly the case with Shankland's experiments. These observations have led Dirac and others to propose that the principle of conservation of energy cannot be retained any longer. Pauli and Jordan, however, consider the energy principle indispensable, and have put forward the suggestion that certain exceedingly small particles are emitted along with β-rays, which account for the necessary energy balance in the disintegration processes. These particles—called " Neutrinos "—are supposed to have no charge and a vanishingly small mass. Many experiments have been carried out to detect them—if they exist—, but so far all attempts have been unsuccessful.

Dirac himself has pointed out that even if the principle of conservation of energy is not possible, that part of quantum mechanics which is non-relativistic, would not be affected. The crucial question itself is, however, by no means settled yet, as Shankland's results have not been confirmed by other experimenters, and throughout the present volume we shall assume that both the energy and momentum principles are valid.

From the energy principle we have therefore in the present case :

$$h\nu = h\nu' + m_0c^2\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right), \tag{1}$$

where m_0 is the mass of the electron and v its velocity of recoil. We use the principle of the conservation of momentum in the two directions along and at right angles to the direction of the incident beam, and get :

$$\frac{h\nu}{c} = \frac{h\nu'}{c}\cos\theta + \frac{m_0v}{\sqrt{1-\frac{v^2}{c^2}}}\cos\theta', \tag{2}$$

$$0 = \frac{h\nu'}{c}\sin\theta - \frac{m_0v}{\sqrt{1-\frac{v^2}{c^2}}}\sin\theta'. \tag{3}$$

Instead of these two equations, we can take the single equation given by the law of composition of two vectors easily deduced from (2) and (3) :

$$\left(\frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}}\right)^2 = \left(\frac{h\nu}{c}\right)^2 + \left(\frac{h\nu'}{c}\right)^2 - 2\,\frac{h\nu}{c}\,\frac{h\nu'}{c}\cos\theta. \tag{4}$$

From the equation (1) we get :

$$\frac{h(\nu-\nu')}{m_0c^2} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1,$$

or, transposing and squaring :

$$\frac{1}{1-\frac{v^2}{c^2}} = \left\{1 + \frac{h(\nu-\nu')}{m_0c^2}\right\}^2$$

$$= 1 + 2\,\frac{h(\nu-\nu')}{m_0c^2} + \frac{h^2(\nu-\nu')^2}{m_0{}^2c^4},$$

so that

$$\frac{1}{1-\frac{v^2}{c^2}} - 1 = 2\,\frac{h(\nu-\nu')}{m_0c^2} + \frac{h^2(\nu-\nu')^2}{m_0{}^2c^4},$$

i.e.,

$$\frac{v^2}{\left(1-\frac{v^2}{c^2}\right)} = \frac{2h(\nu-\nu')}{m_0} + \frac{h^2(\nu-\nu')^2}{m_0{}^2c^2}. \tag{5}$$

Substituting (5) in (4) we get

$$2hm_0(\nu-\nu') + \frac{h^2}{c^2}(\nu-\nu')^2 = \frac{h^2}{c^2}(\nu^2+\nu'^2-2\nu\nu'\cos\theta),$$

or, simplifying and rearranging,

$$2m_0h(\nu-\nu') = \frac{2h^2}{c^2}\,\nu\nu'\,(1-\cos\theta),$$

or, dividing both sides by $2m_0h\nu\nu'$, since $1-\cos\theta = 2\sin^2\frac{\theta}{2}$,

$$\frac{1}{\nu'} - \frac{1}{\nu} = \frac{2h}{m_0c^2}\sin^2\frac{\theta}{2}. \tag{6}$$

If λ, λ' are the wave-lengths of the incident and scattered quanta respectively, then since $\lambda = \frac{c}{\nu}$, $\lambda' = \frac{c}{\nu'}$, we get from (6)

$$\lambda' - \lambda = \frac{2h}{m_0c}\sin^2\frac{\theta}{2} = \frac{h}{m_0c}(1-\cos\theta). \tag{7}$$

This is the fundamental equation of the Compton effect, giving the change in the wave-length of the scattered radiation dependent on the angle of deviation. The value of $\Delta\lambda = \lambda' - \lambda$, found experimentally by Compton is in close agreement with that given by (7).

To find the value of the recoil velocity v of the scattering electron, we put

$$a = \frac{h\nu}{m_0c^2} = \frac{h}{m_0c\lambda_0}\,; \tag{8}$$

then from (6) we have on multiplying both sides by ν:

$$\frac{\nu}{\nu'} - 1 = \frac{2h\nu}{m_0c^2}\sin^2\frac{\theta}{2} = 2a\sin^2\frac{\theta}{2},$$

so that

$$\frac{\nu}{\nu'} = 1 + 2a\sin^2\frac{\theta}{2}. \tag{9}$$

But, from the equation (1), we have on dividing both sides by $h\nu$:

$$1 = \frac{\nu'}{\nu} + \frac{m_0c^2}{h\nu}\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right),$$

or

$$1 - \frac{\nu'}{\nu} = \frac{m_0c^2}{h\nu}\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right),$$

or

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = 1 + \frac{h\nu}{m_0c^2}\left(1 - \frac{\nu'}{\nu}\right),$$

$$= 1 + a\left(1 - \frac{\nu'}{\nu}\right). \tag{10}$$

Substituting (9) in (10), we get:

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = 1 + a\left\{1 - \frac{1}{\left(1 + 2a\sin^2\frac{\theta}{2}\right)}\right\},$$

or

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = 1 + a.\frac{2a\sin^2\frac{\theta}{2}}{1 + 2a\sin^2\frac{\theta}{2}}$$

$$= \frac{1 + 2a(1+a)\sin^2\frac{\theta}{2}}{1 + 2a\sin^2\frac{\theta}{2}}.$$

Therefore, taking inverse square of both sides,

$$1 - \frac{v^2}{c^2} = \frac{\left(1 + 2a\sin^2\frac{\theta}{2}\right)^2}{\left\{1 + 2a(1 + a)\sin^2\frac{\theta}{2}\right\}^2},$$

or

$$\frac{v^2}{c^2} = 1 - \frac{\left(1 + 2a\sin^2\frac{\theta}{2}\right)^2}{\left\{1 + 2a(1 + a)\sin^2\frac{\theta}{2}\right\}^2}$$

$$= \frac{\left\{1 + 2a(1 + a)\sin^2\frac{\theta}{2}\right\}^2 - \left\{1 + 2a\sin^2\frac{\theta}{2}\right\}^2}{\left\{1 + 2a(1 + a)\sin^2\frac{\theta}{2}\right\}^2}$$

$$= \frac{4a^2(2a + a^2)\sin^4\frac{\theta}{2} + 4a^2\sin^2\frac{\theta}{2}}{\left\{1 + 2a(1 + a)\sin^2\frac{\theta}{2}\right\}^2}.$$

Consequently, we get for the velocity v:

$$v = 2ac\sin\frac{\theta}{2}\cdot\frac{\sqrt{(2a + a^2)\sin^2\frac{\theta}{2} + 1}}{1 + 2a(1 + a)\sin^2\frac{\theta}{2}}. \tag{11}$$

This theory of the scattering of an X-ray quantum by an electron shows that the problem is similar to that of the collision of two perfectly elastic smooth balls, for which also, as we know from dynamics, the two laws of conservation of momentum and energy must be obeyed.

CHAPTER V.

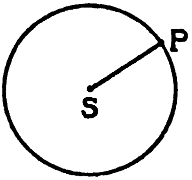
BOHR'S QUANTUM THEORY OF THE ATOM.

5·1. *Classical Theory of the Hydrogen Atom.*

IN sections 3·1 and 3·2, we gave a brief account of Rutherford's researches on the structure of the atom. This picture is, however, still incomplete, as we do not know how the peripheral electrons are arranged in an atom, and how they move. We must apply the classical electrodynamics to examine these questions, but we find that the theory breaks down completely.

Thus, we consider the simplest atom, that of hydrogen, with its own electron revolving round the positively charged nucleus. The mass of the electron is m, its charge $-e$, and the charge on the nucleus is $+e$. It is easily calculated that the electric force between the electron and the proton is about 2×10^{39} times the gravitational force, so that the latter can be neglected.

Suppose, for the sake of simplicity, that the path of the electron is a circle of radius a with its centre at the nucleus.



Let ω be the frequency, *i.e.*, the number of revolutions per second, so that the angular velocity ψ is given by

$$\psi = 2\pi\omega, \tag{1}$$

and the linear velocity at the point P is

$$v = 2\pi a\omega. \tag{2}$$

Now, the equation of motion is

$$m\frac{v^2}{a} = \frac{e^2}{a^2}, \tag{3}$$

so that substituting the value of v from (2) we get

$$\omega^2 = \frac{e^2}{4\pi^2 ma^3}, \quad \omega = \frac{e}{2\pi\sqrt{ma^3}}. \tag{4}$$

Let W be the ionisation potential, *i.e.*, the amount of energy required to remove the electron from the atom. Then if the energy of the ionised atom is taken to be zero, the energy of the normal atom is $-$ W. So that if we denote the total energy of the atom by E, then since the kinetic energy is $\frac{1}{2}mv^2$, and the potential energy is $-\frac{e^2}{a}$, we get

$$\mathrm{E} = -\mathrm{W} = \tfrac{1}{2}mv^2 - \frac{e^2}{a},$$

i.e.,

$$\mathrm{W} = \frac{e^2}{a} - \tfrac{1}{2}mv^2. \tag{5}$$

Substituting in this the value of mv^2 from (3) we get

$$\mathrm{W} = \frac{e^2}{a} - \frac{1}{2}\frac{e^2}{a} = \frac{1}{2}\frac{e^2}{a},$$

thus

$$a = \frac{e^2}{2\mathrm{W}}, \tag{6}$$

and therefore from (4) and (6)

$$\omega^2 = \frac{e^2}{4\pi^2 m}\left(\frac{2\mathrm{W}}{e^2}\right)^3 = \frac{2\mathrm{W}^3}{\pi^2 m e^4}. \tag{7}$$

According to the classical electrodynamics, the electron, being in an accelerated motion, must emit radiation energy, so that E would continuously decrease, *i.e.*, W, being equal to $-$ E, would continuously increase. Therefore, as indicated by the equation (7), ω would continuously increase. Since in a mass of gas, there would be hydrogen atoms in different states having all possible frequencies, and as the frequency ω of the orbit is also the frequency of the corresponding spectral line, the result would be a continuous spectrum, not a line spectrum. But it is well known experimentally that the spectrum of a gaseous element is discontinuous, consisting of definite lines corresponding to definite wave-lengths.

There is yet another difficulty with the classical theory. From equation (6) we see that the radius a of the orbit decreases continuously, and this process will go on until the dimensions of the orbit are of the same order of magnitude as the dimensions of the electrons, or those of the nucleus. The atom would thus be a very unstable system, different from an atomic system occurring in nature, which appears to have absolutely fixed dimensions.

5 · 2. *Bohr's Postulates.*

These considerations led Niels Bohr to the conclusion that the classical electrodynamics is inadequate to describe the behaviour of subatomic phenomena. We have seen in § 4·3, that Max Planck was already forced to adopt this view-point in connection with his theory of black-body radiation. Bohr suggested that Planck's constant h is a universal characteristic of the atom, and must enter into any description of the atomic phenomena. The essential point in Planck's theory of radiation is that, contrary to the classical theory, the energy radiation takes place discontinuously, the amount of energy radiated out from an atomic system of frequency ν being equal to $nh\nu$, where n is an integer.

Generalising Planck's idea, Bohr put forward the following postulates in 1913.

1. The energy radiation is not emitted or absorbed continuously, as assumed in the classical theory, but only during the passage of the system between different "stationary" states.

Translated in terms of Rutherford's atom model, this means that the peripheral electrons cannot move in any arbitrary orbits, varying continuously, but there are certain privileged discrete orbits, on which alone they can move. Each distribution of these electronic orbits represents a stationary state, and associated with each of these stationary states is an energy value. Thus an atomic system has only a discrete set of energy values, or "energy levels" as the stationary states are also called, given by

$$E_1, E_2, \cdots, E_n. \tag{1}$$

The atom is stable in any one of these states, *i.e.*, it can remain indefinitely in a stationary state without radiating energy. Emission or absorption of energy radiation can take place only in a transition of the atom from one stationary state to another.

2. So long as the atomic system is in a particular stationary state E_n, its dynamical equilibrium is governed by the ordinary laws of classical mechanics. But in a transition of the system from the state E_n to another state $E_n{}'$, these laws do not hold any more.

3. The radiation emitted by an atom or molecule during the transition from a stationary state of energy E_2 to another having a lower energy E_1, is homogeneous; the relation between

the frequency ν of the radiation and the total amount ϵ of energy emitted, is given by the formula

$$\epsilon = E_2 - E_1 = h\nu, \tag{2}$$

where h is Planck's constant.

Conversely, the absorption of radiation by an atom or molecule occurs during a transition from a stationary state of energy E_1 to another having a higher energy E_2; and the relation between the frequency ν of radiation and the total amount ϵ of energy absorbed, is given by the same formula (2). The formula (2) is called " Bohr's frequency relation".

4. The different stationary states of a simple system consisting of an electron rotating round a positive nucleus, *i.e.*, for a hydrogen-like atom, were determined by the condition that only such orbits were permissible in which the angular momentum of the electron round the nucleus is an integral multiple of $\frac{h}{2\pi}$, *i.e.*,

$$\text{The angular momentum} = n\,\frac{h}{2\pi}, \tag{3}$$

where n is any one of the integres 1, 2, 3, ····, and where h is again Planck's constant.

The equation (3) for determining the stationary orbits is called " Bohr's quantum condition".

These are the fundamental assumptions applied by Bohr to Rutherford's atom model in his epoch-making paper published in 1913. He was successful in accounting for the rules given by Balmer and Rydberg connecting the frequency of the different lines in the line spectrum of an element.

5 · 3. *Bohr's Theory of the Hydrogen Atom.*

In § 5 · 1 we saw that the classical theory breaks down completely for the atom. However, since in postulate 2 § 5 · 2, Bohr has assumed that the dynamical equilibrium in a stationary state is governed by classical laws, we can take over equations (4) and (6) from § 5 · 1, so that

$$4\pi^2 m a^3 \omega^2 = e^2, \tag{1}$$

and

$$a = \frac{e^2}{2W}, \tag{2}$$

where $W = -E$ is the negative energy or ionisation, in the orbit.

According to Bohr, these two equations are not sufficient for determining the orbit, *i.e.*, the stationary state. We must supplement them by a third equation given by the quantum condition. The linear momentum of the electron is mv, so that the angular momentum, *i.e.*, the moment of momentum about the nucleus, would be mva. Thus from the quantum condition (3) § 5· 2 we get

$$mva = n\frac{h}{2\pi} \quad (n = 1, 2, \cdots). \tag{3}$$

But from (2) § 5· 1, $v = 2\pi a\omega$, so that (3) becomes

$$4\pi^2 m a^2 \omega = nh. \tag{4}$$

Solving (1) and (4) for a and ω, we get

$$a = \frac{h^2 n^2}{4\pi^2 m e^2}, \tag{5}$$

and

$$\omega = \frac{4\pi^2 m e^2}{h^3 n^3}. \tag{6}$$

The possible stationary states, *i.e.*, the possible orbits for the electron, are circles with radii proportional to 1, 4, 9, 16, $\cdots$. The orbit with the smallest radius a_0, given from (5) by putting $n = 1$, is called the "ground orbit",

$$a_0 = \frac{h^2}{4\pi^2 m e^2}. \tag{7}$$

This is the orbit for the "normal" state of the atom, having the minimum energy. All other (larger) orbits are those for the "excited" states. Formula (7) gives the value of a_0 of the order 10^{-8}, which agrees with the value calculated from other methods.

From the equations (2) and (5) we get

$$-E_n = W_n = \frac{e^2}{2a} = \frac{e^2}{2}\cdot\frac{4\pi^2 m e^2}{h^2 n^2} = \frac{2\pi^2 m e^4}{h^2 n^2}, \tag{9}$$

giving the energy E_n of the stationary state. The integer n occurring in the formulæ (5) and (9) characterises the stationary state, and is called the "quantum number of the state".

The frequency ν of the energy radiation emitted during the transition from a state n to another state n' ($n' < n$), is given by Bohr's frequency relation (2) § 5· 2:

$$h\nu = E_n - E_{n'} = W'_n - W_n,$$

or

$$\nu = \frac{E_n - E_{n'}}{h} = \frac{W_n' - W}{h}. \tag{10}$$

Substituting in this the values from (9), we get

$$\nu = \frac{2\pi^2 me^4}{h^3}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right). \tag{11}$$

If c is the velocity of light and λ is the wave-length of the spectral line, then since $\nu = \frac{c}{\lambda}$, we get

$$\frac{1}{\lambda} = \frac{\nu}{c} = \frac{2\pi^2 me^4}{ch^3}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right)$$

$$= \mathrm{R}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right), \tag{12}$$

$$\mathrm{R} = \frac{2\pi^2 me^4}{ch^3}. \tag{13}$$

$\frac{1}{\lambda}$ is called the "wave-number" of the spectral line, and the equation (12) shows that the wave-number of any spectral line is the difference of the wave-numbers of two other lines. This theorem was put forward on purely experimental grounds by W. Ritz in 1908, and is called the "Ritz Combination Principle".

The constant R is the Rydberg constant, and its value 109700 calculated from (13) agrees very closely with the experimentally observed value 109677, providing a strong confirmation of Bohr's theory.

If in the formula (12) we put $n' = 2$, and $n = 3, 4, 5, 6$ successively we get the well-known Balmer lines H_α, H_β, H_γ, H_δ of the hydrogen spectrum. Giving other values to n' and n, we get the other lines of the spectrum.

The same theory holds for any other heavier atom which is ionised so that all except one of its peripheral electrons are removed. The singly ionised helium atom He^+, and the doubly ionised lithium atom Li^{++}, are simple instances of such systems in which only one electron is left revolving round the nucleus. They are usually called hydrogen-like atoms.

Thus if N is the atomic number of the element, the charge on the nucleus is $+ Ne$, so that for the force $\frac{e^2}{r^2}$ between the electron and the nucleus of the hydrogen atom, we must take now $\frac{Ne^2}{r^2}$; consequently in all the previous formulæ e^2 should be replaced by Ne^2. Thus for the equations (5), (9) and (11), we get

$$a = \frac{h^2n^2}{4\pi^2m\,Ne^2}, \tag{14}$$

$$-\,E_n = W_n = \frac{2\pi^2 m N^2 e^4}{h^2 n^2}, \tag{15}$$

$$\nu = \frac{2\pi^2 m N^2 e^4}{h^3}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right). \tag{16}$$

The wave-number of a hydrogen-like atom is given by

$$\frac{1}{\lambda} = \frac{\nu}{c} = \frac{2\pi^2 m e^4}{ch^3} N^2\left(\frac{1}{n'^2} - \frac{1}{n^2}\right)$$

$$= N^2R\left(\frac{1}{n'^2} - \frac{1}{n^2}\right). \tag{17}$$

The spectrum of singly ionised helium is obtained from (17) by writing $N = 2$, and giving various values to n', n. Similarly for doubly ionised lithium we put $N = 3$, and so on.

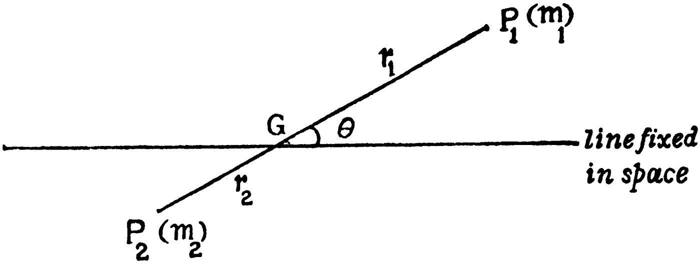
The orbit of the electron in this section has been assumed to be a circle. But from the classical dynamical theory we know that the orbit of a particle under a central force of attraction is an ellipse. However, when calculations are made for the elliptic orbit, it turns out that the values of a, ω and W are exactly the same as those found above for the circular orbit. It is also found that the energy of the orbit depends only on the length of the major axis, being entirely independent of the minor axis, which may have any value. Thus the hypothesis of elliptic orbits brings nothing essentially new in Bohr's theory. It shows only that the emission of a given line can take place in more than one way.

For this reason we shall not reproduce the detailed calculations for the elliptic orbit. In § 5·7 we shall consider the more general problem, taking account of the relativity effect.

5·4. *Motion of the Nucleus.*

In the first presentation of this theory, Bohr considered the mass of the nucleus so large compared to that of the revolving electron, that the nucleus could be taken to be practically at rest. Later he found that results in better agreement with experimental facts can be obtained, if account is taken of the finite mass of the nucleus. This consideration leads us to the problem of two bodies in classical mechanics.

Consider two particles of masses m_1, m_2 moving at a distance r apart under a common attraction F (r) directed along the line joining them. As there is no external force acting on the system the centre of mass G has no acceleration, so that G has a constant velocity in space, and if it is initially at rest it will be always at rest during the subsequent motion of the system.



Suppose that the distances of the two particles P_1 and P_2 from G are r_1 and r_2 respectively, and that the line joining them makes at any time an angle θ with a line fixed in space.

We have from the definition of G, $m_1 r_1 = m_2 r_2$, so that

$$\frac{r_1}{m_2} = \frac{r_2}{m_1} = \frac{r_1 + r_2}{m_1 + m_2} = \frac{r}{m_1 + m_2}, \tag{1}$$

where $r = P_1 P_2$; therefore

$$r_1 = \frac{m_2}{m_1 + m_2} r, \quad r_2 = \frac{m_1}{m_1 + m_2} r. \tag{2}$$

The accelerations of m_1 relative to G along the radius vector G P_1 and perpendicular to it, are

$$\ddot{r}_1 - r\dot{\theta}_1{}^2 \quad \text{and} \; \frac{1}{r_1} \frac{d}{dt} (r_1{}^2 \dot{\theta}) ; \tag{3}$$

but as the accelerations of G are zero, the complete accelerations of m_1 are given by (3). Substituting the value of r_1 from (2) in (3), we get for the complete accelerations of m_1:

$$\frac{m_2}{m_1 + m_2} (\ddot{r} - r\dot{\theta}^2), \; \frac{m_2}{m_1 + m_2} \frac{1}{r} \frac{d}{dt} (r^2 \dot{\theta}). \tag{4}$$

The equations of motion of m_1 are therefore

$$\frac{m_1 m_2}{m_1 + m_2} (\ddot{r} - r\dot{\theta}^2) = - \text{F}, \tag{5}$$

$$\frac{m_1 m_2}{m_1 + m_2} \frac{1}{r} \frac{d}{dt} (r^2 \dot{\theta}) = 0. \tag{6}$$

These equations show that the motion of m_1 relative to m_2 is the same as if m_2 were fixed and m_1 were replaced by $\frac{m_1 m_2}{m_1+m_2}$.

The kinetic energy of the system is

$$T = \tfrac{1}{2}m_1(\dot{r}_1^2 + r_1^2\dot{\theta}^2) + \tfrac{1}{2}m_2(\dot{r}_2^2 + r_2^2\dot{\theta}^2)$$

$$= \tfrac{1}{2}m_1\left\{\frac{m_2^2\dot{r}^2}{(m_1+m_2)^2} + \frac{m_2^2 r^2\dot{\theta}^2}{(m_1+m_2)^2}\right\} + \tfrac{1}{2}m_2\left\{\frac{m_1^2\dot{r}^2}{(m_1+m_2)^2} + \frac{m_1^2 r^2\dot{\theta}^2}{(m_1+m_2)^2}\right\}$$

$$= \tfrac{1}{2}\frac{m_1 m_2}{m_1+m_2}(\dot{r}^2 + r^2\dot{\theta}^2). \qquad (7)$$

This is the same as if m_2 were fixed and m_1 were replaced by $\frac{m_1 m_2}{m_1+m_2}$.

The angular momentum of the system about G is

$$= \tfrac{1}{2}m_1 r_1^2\dot{\theta} + \tfrac{1}{2}m_2 r_2^2\dot{\theta} = \tfrac{1}{2}\frac{m_1 m_2}{m_1+m_2}r^2\dot{\theta}, \qquad (8)$$

and again is the same as if m_2 were fixed and m_1 were replaced by $\frac{m_1 m_2}{m_1+m_2}$.

Thus to allow for the finite mass M_N of the nucleus, we can make the necessary corrections in Bohr's formulæ (14), (15) and (16) of § 3·5 by the mere substitution of $\frac{mM_N}{m+M_N}$ for m, thus giving

$$a = \frac{n^2h^2(m+M_N)}{4\pi^2 mM_N Ne^2}, \qquad (9)$$

$$W = \frac{2\pi^2 mM_N N^2 e^4}{(m+M_N)\,n^2h^2}, \qquad (10)$$

$$\omega = \frac{4\pi^2 mM_N N^2 e^4}{(m+M_N)n^3h^3}. \qquad (11)$$

The frequency of the spectral line due to the transition from a state n to another state n', is given by

$$\nu = \frac{W_{n'} - W_n}{h} = \frac{2\pi^2 mM_N N^2 e^4}{(m+M_N)\,h^3}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right). \qquad (12)$$

It must be remembered that M_N, the mass of the nucleus, is different for atoms of different elements.

As remarked above, the formula (12) gives more correct results than those given by (16) § 5·3. It has also been very useful in clearing doubts about the origin of certain spectral lines left over by the previous formula. The formula (12) can give definite answer as to whether a particular spectral line is due to one element rather than the other.

Thus according to the previous formula, the wave-number of a helium line due to a transition from $2n_1$ to $2n_2$ would be, since $\mathrm{N} = 2$ for helium,

$$\frac{\nu}{c} = \frac{1}{\lambda} = \frac{2\pi^2 m \cdot 4e^4}{ch^3}\left(\frac{1}{4n_2^2} - \frac{1}{4n_1^2}\right)$$

$$= \frac{2\pi\ me^4}{ch^3}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right); \tag{13}$$

and is the same as the wave-number of a hydrogen line due to the transition from n_1 to n_2.

But according to the formula (12), the wave-number of a helium line due to a transition from $2n_1$ to $2n_2$ would be

$$\frac{1}{\lambda} = \frac{2\pi^2 m\mathrm{M}_{\mathrm{He}}\, e^4}{(m + \mathrm{M}_{\mathrm{He}})\, ch^3}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right), \tag{14}$$

while the wave-number of a hydrogen line due to a transition from n_1 to n_2 would be

$$\frac{1}{\lambda} = \frac{2\pi^2 m\mathrm{M}_{\mathrm{H}} e^4}{(m + \mathrm{M}_{\mathrm{H}})\, ch^3}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right). \tag{15}$$

As M_{He}, the mass of the helium nucleus, is different from M_{H}, the mass of the hydrogen nucleus, (14) and (15) would not give the same results.

5·5. *Generalisation of the Quantum Condition.*

In developing Bohr's theory for higher and more complex atomic systems, the first problem would be to generalise the quantum condition (3) § 5·2, *viz.*,

$$\text{The angular momentum} = \frac{nh}{2\pi},\quad (n = 1,\ 2,\ \cdots), \tag{1}$$

which was assumed for a system having a single degree of freedom. In § 5·3 we saw that for a circular orbit of radius a, described with the constant velocity v, (1) became

$$mav = \frac{nh}{2\pi},\quad \text{or } mv\ (2\pi a) = nh,$$

i.e.,

$$mv \int ds = nh,$$

where the integral is taken over the perimeter of the whole circle.

Thus, if σ is the period of a whole revolution, then

$$mv \int_0^\sigma \frac{ds}{dt}\, dt = nh,$$

or, since $\frac{ds}{dt} = v$,

$$mv^2 \int_0^\sigma dt = nh.$$

The kinetic energy T at any point is $T = \frac{1}{2} mv^2$, so that

$$\int_0^\sigma 2T\, dt = nh. \tag{2}$$

But if we use the relation that momentum $p = m\, v$, and velocity $\dot{q} = v$, so that $2T = mv^2 = p\dot{q}$, we can express the quantum condition (2) also in the form

$$\int_0^\sigma p\dot{q}\, dt = nh,$$

or, in the notation of § 1·5, if J is the adiabatically invariant action variable,

$$J = \oint pdq = nh. \tag{3}$$

Thus we see that for a simply periodic system, *i.e.*, one with a single degree of freedom, Bohr's original quantum condition (3) § 5·2 turns out to be equivalent to the statement that the action variable $J = \oint pdq$, which is an adiabatic invariant, should be an integral multiple of h:

$$J = nh\ (n = 1, 2, \cdots). \tag{4}$$

Suppose now that the system is multiply periodic, having s degrees of freedom. We must therefore seek for the quantum conditions for this general case. The guiding principle was supplied by P. Ehrenfest in 1914. Considering the influence of the slowly changing external forces on the atomic system, he came to the conclusion that the quantum conditions are supplied only by equating the adiabatic invariants of the classical theory to nh. This is the "Adiabatic Principle of Ehrenfest".

In § 2·7 we found that for cyclic co-ordinates in the case of non-degenerate systems, we get s action variables defined by

$$J_r = \oint p_r\, dq_r, \quad (r = 1, 2, \cdots, s). \tag{5}$$

Planck, Sommerfeld and Wilson proposed in 1915, that the generalised quantum conditions for systems of many degrees of freedom should be s in number, given by equating the s action variables (2) to integral multiples of h. Thus

$$J_1 = n_1 h,\ J_2 = n_2 h,\ \cdots,\ J_s = n_s h. \tag{6}$$

These conditions are natural generalisations of Bohr's quantum condition (4).

In 1916, Schwarzschild pointed out that although the equations (6) are suitably chosen for non-degenerate systems, for degenerate systems with only u independent frequencies ($u < s$), the imposition of s quantum conditions unduly restricts the number of stationary states. However, for such degenerate systems there are always new action variables J_r', u in number, which can be derived from the J_r by a linear transformation. These J_r' are independent, and Schwarzschild put forward the principle that for degenerate systems the quantum conditions should be

$$J_1' = n_1 h,\ J_2' = n_2 h,\ \cdots,\ J_u' = n_u h. \tag{7}$$

J. M. Burgers proved in 1917 that for non-degenerate systems the J_r, and for degenerate systems the J_r', are adiabatic invariants, so that we see that the conditions (6) or (7), as the case may be, satisfy the adiabatic principle of Ehrenfest, and are therefore the required generalisation of the quantum conditions.

5·6. *The Correspondence Principle.*

The classical theory has been very much useful in a vast domain of the physical world where gross phenomena depending on a large number of atoms are considered. But the theory is insufficient to give us a true picture of a number of other phenomena, such as black-body radiation, atomic collisions, etc. The classical theory will, however, still hold for phenomena which lie on the borderland of these two domains. It is reasonable, therefore, to expect that for such phenomena, the quantum theory should yield no results which differ from those of the classical theory. Thus the fundamental idea of the Correspondence Principle is that the classical theory shall appear as a limiting case of the quantum theory. This implies that as the quantum number n becomes very large, the quantum laws should tend asymptotically to the classical laws.

For, if a periodic system is in a state characterised by a large quantum number n, then the discontinuous variation of the actional variable J, due to the quantum condition $J = nh$, will become quite insignificant. Therefore, the behaviour of the system will be predicted correctly by the classical theory. Thus, the change of energy for a transition of the hydrogen atom from n_1 to n_2 is given by

$$E_{n_1} - E_{n_2} = 2\pi^2 me^4\left(\frac{1}{J_{n_1}^2} - \frac{1}{J_{n_2}^2}\right) = \frac{2\pi^2 me^4}{h^2}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right). \tag{1}$$

If both n_1 and n_2 are large, the value of the right-hand expression is very small, showing that the quantum law demanding a discontinuous change of energy, merges asymptotically for large quantum numbers into the law of a continuous change of energy required by the classical theory.

Now consider an electrically charged particle with charge e, situated at a point (x, y, z). According to classical electrodynamics, the radiation from such a point is determined by its electric moment $\vec{p}$ whose components are given by

$$p_x = ex, \quad p_y = ey, \quad p_z = ez. \tag{2}$$

From equation (6) of § 2·7 we know that if the motion of the particle is simply periodic, the angle variable w, with the period 1, is given by

$$w = \nu_c t + \beta, \tag{3}$$

where ν_c is the frequency in the orbit. Thus each of the co-ordinates x, y, z will be a periodic function of w, and on account of (2), each of the components p_x, p_y, p_z of the electric moment will also be a periodic function of w, and consequently capable of being expanded in a Fourier series of the type

$$\sum_{l=1}^{\infty} C_l \cos 2\pi l w = \sum_{l=1}^{\infty} C_l \cos 2\pi l\,(\nu_c t + \beta). \tag{4}$$

Thus, on the classical theory, the motion of an electrified particle with orbital frequency ν_c gives rise to spectral lines of frequencies $\nu_c, 2\nu_c, 3\nu_c, \ldots$.

Now, in quantum theory, the frequency ν_q of the spectral line due to a transition from n_1 to n_2, is given by

$$\nu_q = \frac{E_{n_1} - E_{n_2}}{h} = (n_1 - n_2)\,\frac{E_{n_1} - E_{n_2}}{J_{n_1} - J_{n_2}} = (n_1 - n_2)\,\frac{\Delta E}{\Delta J}, \tag{5}$$

where we have $J_1 = n_1 h$, $J_2 = n_2 h$.

In § 2·7 we saw that the Hamiltonian H is a function of J only, and since H is the total energy E, we can take E to be a function of J, and that

$$\frac{dE}{dJ} = \frac{dH}{dJ} = \nu_c. \tag{6}$$

Now, if the energy E changes slowly with J, and the quantum numbers n_1 and n_2 are not small compared to the difference $n_1 - n_2$ between them, then we can replace $\frac{\Delta E}{\Delta J}$ in (5) by $\frac{dE}{dJ}$, obtaining from (6):

$$\nu_q = (n_1 - n_2)\,\nu_c = \Delta n \cdot \nu_c. \tag{7}$$

If we put successively $\Delta n = 1, 2, 3, \cdots$ in (7), we get the frequencies ν_q of the spectral lines, according to the quantum theory, by

$$\nu_c, 2\nu_c, 3\nu_c, \cdots, \tag{8}$$

agreeing, for large quantum numbers, with the classical frequencies. It must be remembered that this agreement is only asymptotical, because there can never be any question of an actual identification of the two theories. According to the classical theory, all the frequencies ν_c, $2\nu_c$, $\cdots$, are emitted together by one process, *viz.*, by the description of a single orbit. But according to the quantum theory, each one of the frequencies is emitted in an independent process, *viz.*, by the transitions from orbits of quantum numbers $n + 1$, $n + 2$, $n + 3$, $\cdots$ to that of number n.

This is Bohr's correspondence principle, which says that although the quantum frequency ν_q is, in general, different from the classical frequency $l\nu_c$, in the limiting case of large quantum numbers n the quantum frequencies ν_q for transitions $n + \Delta n$ to n, nearly coincide with the classical frequencies $l\nu_c$, provided Δn is small compared to n.

Bohr further extended his correspondence principle to the intensities and polarisations of the spectral lines. From the classical electrodynamics, we know that the intensity of a spectral line of frequency $l\nu_c$ is proportional to C_l^2, *i.e.*, to the square of the corresponding amplitude in the Fourier expression. The correspondence principle requires that, in the limiting case of large quantum numbers, the classical intensity should agree with the intensity given by the quantum theory. From this fact alone we can draw some inferences about the possibility of the occurrence of transitions. For instance, if the classical intensity of the line with the frequency $l\nu_c$ vanishes, *i.e.*, if the amplitude $C_l = 0$, then according to the classical theory, the spectral line with the frequency $l\nu_c$ is not emitted.

In the quantum theory this signifies that for large n the transition $n + l$ to n does not occur. It is, of course, compatible with the correspondence principle that for small n the transition may be possible. But as an extension of his correspondence principle, Bohr assumed that even when the quantum number n is small, the transition $n + l$ to n is not possible. This can be explained in the following way. The vanishing of a certain coefficient in the Fourier series is a consequence of the symmetry in the geometrical configuration of the system. This symmetry will be effectual in the quantum theory independently of the energy states, and will therefore lead to the impossibility of the corresponding transition.

These considerations provide us with the " selection rules " which explain the anomalies in the occurrence of spectral lines.

The polarisation properties can be determined at the same time as the intensities by a consideration of the components p_x, p_y, p_z, of the electric moment.

5·7. *Relativity Theory of the Fine Structure of Spectral Lines.*

As an application of the foregoing theory we shall consider the problem of the hydrogen atom, taking account of the relativity effect. In § 5·3 we remarked that the elliptic orbits yield nothing more as regards the energy and frequency than the circular orbits considered originally by Bohr. But it is well known on the experimental side that each line of the spectrum is split up into several components, if examined by spectrographs of high resolving power. For instance, the H_α line, which is seen as one with ordinary instruments, is split up into three components when observed through a powerful spectrograph.

Obviously the theory of § 5·3 is unable to account for this "fine structure" of the spectral lines, as it is called. It was Sommerfeld who explained it in 1915, by pointing out that Bohr's theory developed in § 5·3 is still incomplete, inasmuch as it neglects the relativity-effect. In that section we assumed that the mass of the electron is constant throughout, whereas it was proved in § 1·6 (1), that the mass of a moving particle changes with its velocity. For the electron which is revolving rapidly round the nucleus this effect would be appreciable, and ought to be taken into account when the motion is considered. Thus

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{1}$$

where m_0 is the rest mass and v the velocity of the electron. The kinetic energy T is then given by

$$\mathrm{T} = c^2 (m - m_0) = m_0 c^2 \left(\frac{1}{\sqrt{1 - \gamma^2}} - 1\right), \tag{2}$$

where

$$\gamma = \frac{v}{c}.$$

Now consider the elliptic orbit of a hydrogen-like atom of atomic number N. If the polar co-ordinates of the peripheral electron at any point of its orbit are (r, θ), the principle of angular momentum gives one equation of motion

$$mr^2 \frac{d\theta}{dt} = k, \tag{3}$$

where k is a constant. Writing $u = \frac{1}{r}$, we get

$$m \frac{d\theta}{dt} = \frac{k}{r^2} = ku^2,$$

or

$$\frac{d\theta}{dt} = \frac{ku^2}{m}. \tag{4}$$

If (x, y) are the Cartesian co-ordinates of the particle,

$$\begin{aligned}
\frac{d}{dt}(m\dot{x}) &= \frac{d}{d\theta}(m\dot{x})\frac{d\theta}{dt} = \frac{ku^2}{m} \cdot \frac{d}{d\theta}(m\dot{x}) \\
&= \frac{ku^2}{m}\frac{d}{d\theta}\left\{m\frac{dx}{dt}\right\} = \frac{ku^2}{m}\frac{d}{d\theta}\left\{m\frac{dx}{d\theta}\cdot\frac{d\theta}{dt}\right\} \\
&= \frac{ku^2}{m}\frac{d}{d\theta}\left\{ku^2 \cdot \frac{d}{d\theta}(r\cos\theta)\right\} \\
&= \frac{k^2u^2}{m}\frac{d}{d\theta}\left\{u^2\frac{d}{d\theta}\left(\frac{\cos\theta}{u}\right)\right\} \\
&= \frac{k^2u^2}{m}\frac{d}{d\theta}\left\{u^2\left(-\frac{\cos\theta}{u^2}\frac{du}{d\theta} - \frac{\sin\theta}{u}\right)\right\} \\
&= -\frac{k^2u^2}{m}\frac{d}{d\theta}\left\{\frac{du}{d\theta}\cos\theta + u\sin\theta\right\} \\
&= -\frac{k^2u^2}{m}\cos\theta\left(\frac{d^2u}{d\theta^2} + u\right).
\end{aligned}$$

The relativity equation of motion in the direction of the x-axis is

$$\frac{d}{dt}(m\,\dot{x}) = \text{force parallel to the } x\text{-axis.}$$

Therefore, since the force on the electron is $-\frac{\mathrm{N}e^2}{r^2}$ towards the nucleus, we get

$$-\frac{k^2u^2}{m}\cos\theta\left(\frac{d^2u}{d\theta^2}+u\right) = -\frac{\mathrm{N}e^2}{r^2}\cos\theta = -\mathrm{N}e^2u^2\cos\theta,$$

or

$$\frac{d^2u}{d\theta^2}+u = \frac{\mathrm{N}e^2m}{k^2} = \frac{\mathrm{N}e^2m_0}{k^2}\,\frac{1}{\sqrt{1-\gamma^2}}. \tag{5}$$

If V is the potential energy, E the total energy, then

$$\mathrm{T}+\mathrm{V} = \text{constant} = \mathrm{E},$$

so that, since $\mathrm{V} = -\frac{\mathrm{N}e^2}{r} = -\mathrm{N}e^2u$, we get from (2)

$$m_0c^2\left(\frac{1}{\sqrt{1-\gamma^2}}-1\right)-\mathrm{N}e^2u = \mathrm{E}. \tag{6}$$

To get the differential equation of the orbit, we must substitute the value of $\sqrt{1-\gamma^2}$ from (6) in (5). Now from (6) we get

$$\frac{1}{\sqrt{1-\gamma^2}} = \frac{\mathrm{E}+\mathrm{N}e^2u}{m_0c^2}+1\,;$$

therefore substituting this in (5) we get

$$\frac{d^2u}{d\theta^2}+u = \frac{\mathrm{N}e^2m_0}{k^2}\left(\frac{\mathrm{E}+\mathrm{N}e^2u}{m_0c^2}+1\right),$$

or

$$\frac{d^2u}{d\theta^2}+u\left(1-\frac{\mathrm{N}^2e^4}{k^2c^2}\right) = \frac{\mathrm{E}+m_0c^2}{k^2c^2}\,\mathrm{N}e^2.$$

Writing

$$\left.\begin{aligned}\lambda^2 &= 1-\frac{\mathrm{N}^2e^4}{k^2c^2},\\ \mathrm{A} &= \frac{\mathrm{E}+m_0c^2}{k^2c^2}\,\mathrm{N}e^2,\end{aligned}\right\} \tag{7}$$

we get finally

$$\frac{d^2u}{d\theta^2}+\lambda^2u = \mathrm{A}, \tag{8}$$

which is the differential equation of the orbit. We can write the equation (8) in the form

$$\frac{d^2}{d\theta^2}\left(u-\frac{\mathrm{A}}{\lambda^2}\right)+\lambda^2\left(u-\frac{\mathrm{A}}{\lambda^2}\right) = 0,$$

of which the solution is immediately recognised to be

$$u - \frac{A}{\lambda^2} = B\cos(\lambda\theta + \beta)$$

or by choosing the initial line so that the initial phase $\beta = 0$, we get

$$u = B\cos\lambda\theta + \frac{A}{\lambda^2}.$$

Now write

$$l = \frac{\lambda^2}{A}; \quad \frac{B\lambda^2}{A} = \epsilon; \tag{9}$$

Then the equation of the orbit becomes

$$\frac{l}{r} = 1 + \epsilon\cos\lambda\theta. \tag{10}$$

In the limiting case when $c \rightarrow \infty$, *i.e.*, $\lambda \rightarrow 1$, the equation (10) reduces to that of an ordinary conic.

The minimum value of r is $\frac{l}{1+\epsilon}$ when $\cos\lambda\theta = 1$, *i.e.*, when $\lambda\theta = 0, 2\pi, 4\pi, \cdots\cdot$; and the maximum value of r is $\frac{l}{1-\epsilon}$ when $\cos\lambda\theta = -1$, *i.e.*, when $\lambda\theta = \pi, 3\pi, 5\pi, \cdots\cdot$.

Thus the apses of the orbits are $r = \frac{l}{1+\epsilon}$ and $r = \frac{l}{1-\epsilon}$ for the above values of $\lambda\theta$, and therefore the angle between the successive apse lines is given by

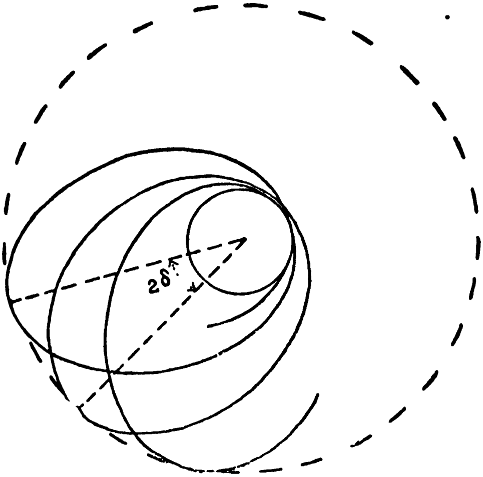
$$\frac{\pi}{\lambda} = \pi\left(1 - \frac{N^2e^4}{k^2c^2}\right)^{-\frac{1}{2}} = \pi\left(1 + \frac{N^2e^4}{2k^2c^2}\right) = \pi + \delta,$$

where $\delta = \frac{\pi}{2}\frac{N^2e^4}{k^2c^2}$. The small angle between successive lines to the perihelion is therefore

$$2\cdot\delta = \frac{\pi N^2e^4}{k^2c^2}. \tag{11}$$

The orbit is therefore an ellipse rotating in such a way as to produce this advance of the perihelion.

The orbit is of a rosette type whose form is that shown in the accompanying figure. This shows that while the electron is going round the nucleus in an ellipse, the ellipse itself is revolving slowly.



If ω is the frequency of the electron in the ellipse, then the frequency of the rotation of the line is ω', given by

$$\omega' = \frac{\omega\delta}{\pi} = \frac{\omega N^2 e^4}{2k^2c^2}. \tag{12}$$

Thus the atomic system has a doubly periodic motion having two frequencies ω, ω'. The fine structure of the spectral lines is accounted for by the introduction of this new frequency ω'.

To give a quantitative account of the resolution of the spectral lines, we must calculate the energy, and the action and angle variables.

If (r, θ) are the polar co-ordinates of the electron, then

$$H = T + V = m_0c^2\left(\frac{1}{\sqrt{1-\gamma^2}} - 1\right) - \frac{Ne^2}{r},$$

where

$$\gamma^2 = \frac{v^2}{c^2} = \frac{\dot{r}^2 + r^2\dot{\theta}^2}{c^2}.$$

If p_1 and p_2 are momenta conjugate to the co-ordinates r, θ respectively, then

$$p_1 = m\dot{r} = \frac{m_0\dot{r}}{\sqrt{1-\gamma^2}}, \quad p_2 = mr^2\dot{\theta} = \frac{m_0r^2\dot{\theta}}{\sqrt{1-\gamma^2}}.$$

We must eliminate $\dot{r}$, $\dot{\theta}$ from H, and express it as a function of r, θ, p_1 and p_2 only, as required by the dynamical theory of § 2·5. We have

$$p_1^2 + \frac{p_2^2}{r^2} = \frac{m_0^2}{1-\gamma^2}(\dot{r}^2 + r^2\dot{\theta}^2) = \frac{m_0^2\gamma^2c^2}{1-\gamma^2}.$$

Therefore

$$m_0^2c^2 + p_1^2 + \frac{p_2^2}{r^2} = m_0^2c^2 + \frac{m_0^2c^2\gamma^2}{1-\gamma^2} = \frac{m_0^2c^2}{1-\gamma^2},$$

so that

$$\frac{m_0c^2}{\sqrt{1-\gamma^2}} = c\sqrt{m_0^2c^2 + p_1^2 + \frac{p_2^2}{r^2}}.$$

Substituting this in H, we get

$$H = c\sqrt{m_0^2c^2 + p_1^2 + \frac{p_2^2}{r^2}} - m_0c^2 - \frac{Ne^2}{r}. \tag{13}$$

Writing $p_1 = \frac{\partial S}{\partial r}$, $p_2 = \frac{\partial S}{\partial \theta}$, we get the Hamilton-Jacobi equation :

$$c\left\{m_0^2c^2 + \left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2\right\}^{\frac{1}{2}} - m_0c^2 - \frac{Ne^2}{r} = \alpha_1, \tag{14}$$

where α_1 is the total energy. By an easy calculation it can be shown that this equation reduces to

$$\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 = \frac{1}{c^2}\left(\alpha_1 + \frac{Ne^2}{r}\right)\left(\alpha_1 + \frac{Ne^2}{r} + 2m_0c^2\right). \tag{15}$$

But from Hamilton's canonical equation $\frac{\partial H}{\partial \theta} = -\dot{p}_2$, and since H does not contain θ, as seen from (13), we, get $\dot{p}_2 = 0$, therefore

$$\frac{\partial S}{\partial \theta} = p_2 = \text{constant} = \alpha_2 \text{ (say)}. \tag{16}$$

From (15) and (16) we have then

$$\left(\frac{\partial S}{\partial r}\right)^2 = -\frac{\alpha_2^2}{r^2} + \frac{1}{c^2}\left(\alpha_1 + \frac{Nc^2}{r}\right)\left(\alpha_1 + \frac{Ne^2}{r} + 2m_0c^2\right)$$

$$= \left(2m_0 + \frac{\alpha_1}{c_2}\right)\alpha_1 + 2Ne^2\left(m_0 + \frac{\alpha_1}{c^2}\right)\frac{1}{r} - \left(\alpha_2^2 - \frac{N^2e^4}{c^2}\right)\frac{1}{r^2};$$

so that

$$\left(\frac{\partial S}{\partial r}\right)^2 = -A + \frac{2B}{r} - \frac{C}{r^2}, \tag{17}$$

where

$$\left.\begin{aligned} A &= -\alpha_1\left(2m_0 + \frac{\alpha_1}{c^2}\right), \\ B &= Ne^2\left(m_0 + \frac{\alpha_1}{c^2}\right), \\ C &= \alpha_2^2 - \frac{N^2e^4}{c^2}. \end{aligned}\right\} \tag{18}$$

Now we must find the action variables J_1, J_2. Thus

$$J_1 = \oint p_1\,dr = \oint \frac{\partial S}{\partial r}\,dr = \oint \sqrt{-A + \frac{2B}{r} - \frac{C}{r^2}}\,dr, \qquad (18')$$

$$= \oint \frac{dr}{r}\sqrt{-Ar^2 + 2Br - C} = \frac{dr}{r}\sqrt{A(\alpha - r)(r - \beta)},$$

where α, β are the roots of the equation $+ r^2 - \frac{2B}{A}r + \frac{C}{A} = 0$, *i.e.*, where $\alpha\beta = \frac{C}{A}$, $\alpha + \beta = \frac{2B}{A}$. Thus β is the minimum value of r, and α the maximum value. Since the integration is extended over a whole period, *i.e.*, from $r_{\text{min.}}$ to $r_{\text{max.}}$ and back to $r_{\text{min.}}$, we get

$$J_1 = 2\sqrt{A}\cdot\int_\beta^\alpha \frac{dr}{r}\sqrt{(\alpha - r)(r - \beta)}.$$

Make the transformation $r = \alpha\sin^2\phi + \beta\cos^2\phi$, then

$$J_1 = 2\sqrt{A}\cdot\int_0^{\frac{\pi}{2}} \frac{(\alpha - \beta)^2\, 2\sin^2\phi\cos^2\phi}{\alpha\sin^2\phi + \beta\cos^2\phi}\,d\phi,$$

$$= 4\sqrt{A}\,(\alpha - \beta)^2\int_0^{\frac{\pi}{2}} \frac{\sin^2\phi - \sin^4\phi}{(\alpha - \beta)\sin^2\phi + \beta}\,d\phi,$$

$$= 4\sqrt{A}\,(\alpha - \beta)^2$$

$$\times\int_0^{\frac{\pi}{2}}\left\{-\frac{\sin^2\phi}{(\alpha - \beta)} + \frac{\alpha}{(\alpha - \beta)^2} - \frac{\dfrac{\alpha\beta}{(\alpha - \beta)^2}}{(\alpha - \beta)\sin^2\phi + \beta}\right\}d\phi,$$

$$= 4\sqrt{A}\,(\alpha - \beta)^2 \times$$

$$\times\left\{\frac{-\pi}{4(\alpha - \beta)} + \frac{\pi\alpha}{2(\alpha - \beta)^2} - \frac{\alpha\beta}{(\alpha - \beta)^2}\int_0^{\frac{\pi}{2}}\frac{d\phi}{\alpha\sin^2\phi + \beta\cos^2\phi}\right\}$$

$$= 4\sqrt{A}\,(\alpha - \beta)^2\left\{\frac{-\pi}{4(\alpha - \beta)} + \frac{\pi\alpha}{2(\alpha - \beta)^2} - \frac{\alpha\beta}{(\alpha - \beta)^2}\cdot\frac{\pi}{2\sqrt{\alpha\beta}}\right\},$$

$$= \pi\sqrt{A}\,(\alpha + \beta - 2\sqrt{\alpha\beta}),$$

$$= \pi\sqrt{A}\left(\frac{2B}{A} - 2\sqrt{\frac{C}{A}}\right) = 2\pi\left(\frac{B}{\sqrt{A}} - \sqrt{C}\right). \qquad (18'')$$

Substituting the values of A, B, C from (18) we get

$$J_1 = \frac{2\pi N e^2 \left(m_0 + \frac{\alpha_1}{c^2}\right)}{\sqrt{-\alpha_1\left(2m_0 + \frac{\alpha_1}{c^2}\right)}} - 2\pi \sqrt{\alpha_2^2 - \frac{N^2e^4}{c^2}}. \tag{19}$$

Similarly on account of (16) we have

$$J_2 = \oint p_2 d\theta = \int_0^{2\pi} \frac{\partial S}{\partial \theta} d\theta = \int_0^{2\pi} \alpha_2 d\theta = 2\pi\alpha_2. \tag{20}$$

We must determine from (19) and (20) the value of the total energy α_1 in terms of J_1 and J_2.

Let

$$a^2 = \frac{4\pi^2 e^4}{c^2h^2}, \quad a = \frac{2\pi e^2}{ch}, \tag{21}$$

where h is Planck's constant. a is called "Sommerfeld's fine structure constant," and it is found that $a^2 = 5{\cdot}31 \times 10^{-5}$.

Now we have from (20) and (21):

$$2\pi \sqrt{\alpha_2^2 - \frac{N^2e^4}{c^2}} = \sqrt{4\pi^2\alpha_2^2 - \frac{4\pi^2N^2e^4}{c^2}} = \sqrt{J_2^2 - h^2a^2N^2}.$$

Substituting this in (19) we get

$$J_1 + \sqrt{J_2^2 - h^2a^2N^2} = \frac{chaN\left(m_0 + \frac{\alpha_1}{c^2}\right)}{\sqrt{-\alpha_1\left(2m_0 + \frac{\alpha_1}{c^2}\right)}}$$

$$= \frac{haN\left(m_0 + \frac{\alpha_1}{c^2}\right)}{\sqrt{m_0^2 - \left(m_0 + \frac{\alpha_1}{c^2}\right)^2}}.$$

But

$$\alpha_1 = E = -W, \text{ so that } W = -\alpha_1;$$

therefore

$$J_1 + \sqrt{J_2^2 - h^2a^2N^2} = \frac{haN\left(m_0 - \frac{W}{c^2}\right)}{\sqrt{m_0^2 - \left(m_0 - \frac{W}{c^2}\right)^2}}.$$

Let

$$u = 1 - \frac{W}{m_0c^2}, \tag{21a}$$

then

$$J_1 + \sqrt{J_2^2 - h^2\alpha^2N^2} = \frac{h\alpha Nu}{\sqrt{1-u^2}},$$

or

$$\frac{1-u^2}{u^2} = \frac{h^2\alpha^2N^2}{(J_1 + \sqrt{J_2^2 - h^2\alpha^2N^2})^2},$$

or

$$\frac{1}{u^2} = 1 + \frac{h^2\alpha^2N^2}{\left\{J_1 + J_2\left(1 - \frac{h^2\alpha^2N^2}{J_2^2}\right)^{\frac{1}{2}}\right\}^2}. \tag{21 b}$$

Since from (21), α is a very small quantity, we neglect powers of α greater than 4, so that

$$\begin{aligned}
\frac{1}{u^2} &= 1 + h^2\alpha^2N^2\left\{J_1 + J_2\left(1 - \frac{h^2\alpha^2N^2}{2J_2^2}\right)\right\}^{-2},\\
&= 1 + h^2\alpha^2N^2\left\{(J_1 + J_2) - \frac{h^2\alpha^2N^2}{2J_2}\right\}^{-2},\\
&= 1 + \frac{h^2\alpha^2N^2}{(J_1 + J_2)^2}\left\{1 - \frac{h^2\alpha^2N^2}{2J_2(J_1 + J_2)}\right\}^{-2},\\
&= 1 + \frac{h^2\alpha^2N^2}{(J_1 + J_2)^2}\left\{1 + \frac{h^2\alpha^2N^2}{J_2(J_1 + J_2)}\right\},\\
&= 1 + \frac{h^2\alpha^2N^2}{(J_1 + J_2)^2} + \frac{h^4\alpha^4N^4}{J_2(J_1 + J_2)^3}.
\end{aligned}$$

Therefore

$$\begin{aligned}
u &= \left\{1 + \frac{h^2\alpha^2N^2}{(J_1 + J_2)^2} + \frac{h^4\alpha^4N^4}{J_2(J_1 + J_2)^3}\right\}^{-\frac{1}{2}},\\
&= 1 - \frac{h^2\alpha^2N^2}{2(J_1 + J_2)^2} - \frac{h^4\alpha^4N^4}{2J_2(J_1 + J_2)^3} + \frac{3}{8}\,\frac{h^4\alpha^4N^4}{(J_1 + J_2)^4},\\
&= 1 - \frac{h^2\alpha^2N^2}{2(J_1 + J_2)^2} - \frac{h^4\alpha^4N^4(J_2 + 4J_1)}{8J_2(J_1 + J_2)^4}.
\end{aligned}$$

Substituting the value of u, we get

$$1 - \frac{W}{m_0c^2} = 1 - \frac{h^2\alpha^2N^2}{2(J_1 + J_2)^2} - \frac{h^4\alpha^4N^4(J_2 + 4J_1)}{8J_2(J_1 + J_2)^4}.$$

Therefore, we get finally for the negative energy

$$\begin{aligned}
W &= \frac{m_0c^2h^2\alpha^2N^2}{2(J_1 + J_2)^2}\left\{1 + \frac{h^2\alpha^2(J_2 + 4J_1)N^2}{4J_2(J_1 + J_2)^2}\right\}\\
&= \frac{m_0c^2h^2\alpha^2N^2}{2(J_1 + J_2)^2}\left\{1 + \frac{h^2\alpha^2N^2}{(J_1 + J_2)^2}\left(\frac{1}{4} + \frac{J_1}{J_2}\right)\right\}
\end{aligned} \tag{22}$$

The quantum conditions are

$$J_1 = n_1 h, \qquad J_2 = n_2 h,$$

where n_1 and n_2 are integers, so (22) becomes

$$W = \frac{m_0 c^2 h^2 a^2 N^2}{2\,(n_1 h + n_2 h)^2}\left\{1 + \frac{h^2 a^2 N^2}{(n_1 h + n_2 h)^2}\left(\tfrac{1}{4} + \frac{n_1 h}{n_2 h}\right)\right\}.$$

We take new integers n and l instead of n_1, n_2, satisfying

$$n = n_1 + n_2, \qquad n_2 = l + 1, \tag{23}$$

so that

$$l + 1 \leqslant n, \qquad i.e., l \leqslant n - 1. \tag{24}$$

The inequality (24) shows that for any given integer n, l can take only the following n values

$$0, 1, 2, \cdots (n-2), (n-1). \tag{25}$$

It is customary in spectroscopy to call n the " principle " quantum number, and l the " azimuthal " quantum number.

Thus on substituting the value of a^2, n_1 and n_2, we get

$$\begin{aligned} W &= \frac{m_0 c^2 a^2 N^2}{2\,(n_1 + n_2)^2}\left\{1 + \frac{a^2 N^2}{(n_1 + n_2)^2}\left(\tfrac{1}{4} + \frac{n_1}{n_2}\right)\right\} \\ &= \frac{2\pi^2 m_0 N^2 e^4}{h^2 n^2}\left\{1 + \frac{a^2 N^2}{n^2}\left(\tfrac{1}{4} + \frac{n - n_2}{n_2}\right)\right\} \\ &= \frac{2\pi^2 m_0 N^2 e^4}{h^2 n^2}\left\{1 + \frac{a^2 N^2}{n^2}\left(\frac{n}{l+1} - \tfrac{3}{4}\right)\right\}. \end{aligned} \tag{26}$$

Each stationary state is therefore characterised by two integers n and l, and is denoted as the n_l state. In the limiting case when $c \to \infty$, $a \to 0$ and (26) reduces to the old formula of Bohr's theory.

To be able to apply (26) to calculate frequencies of spectral lines due to transitions, we must have selection rules which would be provided by the correspondence principle.

The transformation function S is obtained by the integration

$$\begin{aligned} S &= \int \frac{\partial S}{\partial r}\, dr + \int \frac{\partial S}{\partial \theta}\, d\theta \\ &= \int \frac{\partial S}{\partial r}\, dr + a_2 \theta. \end{aligned}$$

Now $\frac{\partial S}{\partial r}$ is a function of r, a_1 and a_2, and on account of (20) and (22), a_1 and a_2 are functions of J_1, J_2. Hence $\frac{\partial S}{\partial r}$ is a function of r, J_1, J_2, and so we have

$$S = f(r, J_1, J_2) + \frac{J_2\theta}{2\pi}. \tag{27}$$

If w_1, w_2 are angle variables conjugate to J_1, J_2 respectively, then we get

$$\begin{aligned} w_1 &= \frac{\partial S}{\partial J_1} = f_{J_1}'(r, J_1, J_2), \\ w_2 &= \frac{\partial S}{\partial J_2} = f_{J_2}'(r, J_1, J_2) + \frac{\theta}{2\pi}. \end{aligned} \tag{28}$$

Solving equations (28) for r and θ, we get

$$\begin{aligned} r &= \phi(w_1, J_1, J_2), \\ \theta &= 2\pi w_2 + \psi(w_1, J_1, J_2). \end{aligned} \tag{29}$$

Therefore, if x and y are the Cartesian co-ordinates and $i = \sqrt{-1}$,

$$\begin{aligned} x + iy = re^{i\theta} &= \phi(w_1, J_1, J_2)\, e^{2\pi i w_2 + i\psi(w_1, J_1, J_2)} \\ &= \Phi(w_1, J_1, J_2)\, e^{2\pi i w_2}. \end{aligned}$$

Writing $\Phi(w_1, J_1, J_2)$ as a Fourier series $\sum_k C_k\, e^{2\pi i k w_1}$, where C_k is a function of J_1, J_2, we have

$$\begin{aligned} x + iy &= \sum_k C_k\, e^{2\pi i (kw_1 + w_2)} \\ &= \sum_k D_k\, e^{2\pi i(k\nu_1 + \nu_2)t}, \end{aligned}$$

where ν_1 and ν_2 are frequencies given according to

$$w_1 = \nu_1 t + \beta_1, \qquad w_2 = \nu_2 t + \beta_2.$$

Thus we see that x, y, and therefore the components p_x, p_y of the electric moment are Fourier series of the type

$$\sum_k C_k \cos\{2\pi(k\nu_1 + \nu_2)t\}. \tag{30}$$

The quantum number n_1 corresponds to ν_1, and n_2 corresponds to ν_2. The correspondence principle now shows that $|\Delta n_1|$ may have any value, but $|\Delta n_2| = 1$, *i.e.*, $n = n_1 + n_2$ may change by any amount, but $l = n_2 - 1$ can only change at the most by unity. This appears from the fact that in (30) the coefficient of ν_1 is k which can have any value, whereas the coefficient of ν_2 is unity. Consequently

$$\Delta l = +1, \ -1, \ 0. \tag{30 a}$$

As an illustration, we shall apply this result to the calculation of the fine structure of the H_α line of the Balmer series. For Hydrogen, N = 1, and the frequency ν of the spectral line due to the transition $n \rightarrow n'$, $l \rightarrow l'$ is given by

$$\nu = \frac{E - E'}{h} = \frac{W' - W}{h}$$
$$= \frac{2\pi^2 m_0 e^4}{h^3 n'^3}\left\{1 + \frac{a^2}{n'^2}\left(\frac{n'}{l' + 1} - \tfrac{3}{4}\right)\right\}$$
$$- \frac{2\pi^2 m_0 e^4}{h^3 n^3}\left\{1 + \frac{a^2}{n^2}\left(\frac{n}{l + 1} - \tfrac{3}{4}\right)\right\}.$$

If ν_0 is the frequency of the original line without relativity effect, and R is the Rydberg constant, then

$$\nu - \nu_0 = Rca^2\left\{\frac{1}{n'^4}\left(\frac{n'}{l' + 1} - \tfrac{3}{4}\right) - \frac{1}{n^4}\left(\frac{n}{l + 1} - \tfrac{3}{4}\right)\right\}. \tag{31}$$

Thus the single line of frequency ν_0 splits up into several lines of frequencies given by (31).

For the line H_α, $n = 3$, $n' = 2$. From (23) we see therefore that $l \leqslant 3 - 1$, $l' \leqslant 2 - 1$, *i.e.*, l can have the three values 0, 1, 2, and l' can have the two values 0, 1. From the correspondence principle we know that l can change only by unity, *i.e.*, $l' - l = +1$ or -1, so that we have the following possible pairs of values for l and l':

$$\left.\begin{matrix} l = 0 \\ l' = 1 \end{matrix}\right\} \quad \left.\begin{matrix} l = 1 \\ l' = 0 \end{matrix}\right\} \quad \left.\begin{matrix} l = 2 \\ l' = 1 \end{matrix}\right\}. \tag{32}$$

Substituting these pairs of values in (31) we get the three resolved lines of H_α. The formula is found to give values in close agreement with experiment.

The fine structure of the lines of any other element can be calculated similarly from (31).

5·8. *Development of Bohr's Theory.*

During the years 1915–22, Bohr's theory was applied with success to a number of atomic phenomena, such as the normal Zeeman-effect produced by an external magnetic field, and the

Stark-effect produced by an external electric field. The Hamilton-Jacobi theory of dynamics, in conjunction with Bohr's frequency relation and quantum conditions, gave the energy in the stationary states with sufficient approximation, while an application of the correspondence principle supplied the necessary selection and polarisation rules.

We shall consider these problems according to modern quantum mechanics, in subsequent chapters. But they were explained more or less satisfactorily even on Bohr's theory.

Thus, it was found that when we bring an external electric or magnetic field in a definite direction, we have to introduce a new quantum number m, connected with the angle ϕ of the electron round the axis of the field. For given values of the principle quantum number n and the azimuthal quantum number l, the new quantum number m can assume only the following $2l + 1$ values :

$$m = -l, -(l-1), \cdots, -1, 0, +1, \cdots, (l+1), l. \tag{1}$$

It was also found that only those transitions could take place for which m changes at the most by unity, *i.e.*,

$$\Delta m = -1, 0, +1. \tag{2}$$

5·8 (1). *Atomic Magnetism. The Magnetic Moment of the Atom.*

Consider an electron moving round in a small circle. According to the classical electron theory, an electron in motion generates a magnetic field, identical in all respects with the field of a small magnet placed at the centre, so that the axis of the magnet is perpendicular to the plane of the electronic orbit. If J is the angular momentum of the electron in the orbit, the magnetic moment M of the equivalent bar magnet is given by

$$\mathrm{M} = -\frac{e\mathrm{J}}{2m_0c}, \tag{1}$$

where e is the charge and m_0 the mass of the electron.

To prove this theorem, suppose that r, θ are the polar co-ordinates of the electron, A the area of the orbit, and σ the period. Then $\frac{d\mathrm{A}}{dt}$ would be the rate of description of area, whose value is known to be :

$$\frac{d\mathrm{A}}{dt} = \tfrac{1}{2}\, r^2 \frac{d\theta}{dt}.$$

And since $J = m_0 r^2 \dfrac{d\theta}{dt}$, we get

$$\frac{dA}{dt} = \frac{J}{2m_0}. \tag{2}$$

Integrating this over a whole period, and remembering that J is constant for a circular orbit, we have

$$A = \int_0^\sigma \frac{J}{2m_0}\, dt = \frac{\sigma J}{2m_0}. \tag{3}$$

If the frequency of the electron in the orbit is $\omega = 1/\sigma$, then the electron passes a given point on its orbit ω times per second, and therefore, is equivalent to a current $i = -\dfrac{e}{c}\,\omega$. Now, according to the electrodynamic theory, a current i flowing in a closed circuit of area A, gives rise to the same field at a great distance as a bar magnet of moment

$$M = i\,A, \tag{4}$$

situated at the centre of the orbit, with its axis perpendicular to the plane of the orbit. Since $i = -\dfrac{e}{c}\,\omega$, we have from (4)

$$M = -\frac{e}{c}\,\omega A. \tag{5}$$

But from equation (3) we get

$$\omega A = \frac{A}{\sigma} = \frac{J}{2m_0},$$

therefore

$$M = -\frac{Je}{2m_0 c}. \tag{6}$$

The direction of M is opposed to that of J, because the charge of the electron is negative. Thus, we see that the value of the magnetic moment of the equivalent bar magnet is that given by (1). For an atom having several electrons, we have simply to sum up the contributions of each one of them to get an expression of the same form as (6), but J would then denote the angular momentum of the whole system.

Thus we see that the orbital motion of the electrons in an atom produces a magnetic field, which comes into evidence in the Zeeman-effect.

Now, from Bohr's quantum theory of the circular orbit, we know that the angular momentum is an integral multiple of $\frac{h}{2\pi}$, so that we have

$$J = m\frac{h}{2\pi}, \qquad (m = 1, 2, \cdots). \tag{7}$$

Substituting this in (1) we get,

$$M = m\frac{eh}{4\pi m_0 c}, \qquad (m \text{ integer}). \tag{8}$$

For the lowest orbit, characterised by the quantum number $m = 1$, the magnetic moment μ of the atom is given by

$$\mu = \frac{eh}{4\pi m_0 c} = 0\cdot922 \times 10^{-20} \text{ gauss centimetres.} \tag{9}$$

μ is called the " Bohr-Magneton ", and it is taken as the unit of atomic magnetism. For an atom whose Zeeman-effect is normal, the magnetic moment measured in magnetons is $m\mu$, where m is any integer.

The existence of atomic magnetism, confirming the value of M given above, was demonstrated experimentally by W. Gerlach and O. Stern in 1921, and in subsequent years. The first experiment consisted in passing the atoms of silver vapour through a non-uniform magnetic field, and measuring the deviation from their otherwise rectlinear paths by intercepting them on a screen. From this deviation the value of the magnetic moment can be calculated, which agrees closely with the value given in (8).

5·8 (2). *The Spectra of Heavier Atoms.*

From the atom-model proposed by Rutherford we know that a neutral atom of any element other than hydrogen consists of a positively charged nucleus, around which several electrons are rotating. Naturally, such an atom must have a multiplicity of stationary states, because there are several possibilities for the revolving electrons to arrange themselves in different orbits. The spectra of these atoms are therefore more complicated than the spectrum of the hydrogen atom. Their characteristic feature is that they can be grouped in sequences of lines called " series ". The four main series are the " principal ", " sharp ", " diffuse " and " fundamental " series.

In § 5·7, we saw that if we take account of the relativity effect, the energy E of the atom in any stationary state is characterised by two quantum numbers n and l, where n is the principal quantum number, and l a second quantum number associated with the series to which the term belongs. Thus $W = -E$ is a function of n and l, which we can denote by $W(n, l)$. We know also that $l < n$, and that l can change by 1 only.

It is customary in spectroscopy to denote

(1) $W(n, 0)$, $(l = 0)$, $(n = 1, 2, 3, \cdots)$ as S terms,

(2) $W(n, 1)$, $(l = 1)$, $(n = 2, 3, 4, \cdots)$ as P terms,

(3) $W(n, 2)$, $(l = 2)$, $(n = 3, 4, 5, 6, \cdots)$ as D terms,

(4) $W(n, 3)$, $(l = 3)$, $(n = 4, 5, 6, \cdots)$ as F terms.

A line due to a transition $n = 5$, $l = 2 \rightarrow n' = 2$, $l' = 1$ is denoted by the notation ${}^5D \rightarrow {}^2P$.

The four series mentioned above are given by the transitions

$n\,P \rightarrow 1\,S$ $(n = 2, 3, 4, \cdots)$: Principal series ;

$n\,S \rightarrow 2\,P$ $(n = 2, 3, 4, \cdots)$: Sharp series ;

$n\,D \rightarrow 2\,P$ $(n = 3, 4, 5, \cdots)$: Diffuse series ;

$n\,F \rightarrow 3\,D$ $(n = 4, 5, 6, \cdots)$: Fundamental series.

The series lines, when observed through spectrographs of high resolving power, are found, in general, to be multiple in structure ; they are therefore called "multiplets". Thus it is found that, in the case of the alkali metals, the S terms remain single, but the P, D and F terms each split up into two, giving rise to doublets. Thus the alkalies Na, K, etc., have singlets and doublets. Similarly the alkaline earths Mg, Ca, etc., have terms which are singlets and triplets. To take a concrete example, in the spectrum of sodium, the first line of the principal series, the yellow D-line is found, when examined by spectrographs of high resolving power, to be really two lines separated by 6 Å.

To explain this multiplet structure, Sommerfeld and Landé proposed that two quantum numbers n, l are not sufficient to describe the terms, but that a third quantum number 'm' should also be introduced corresponding to the magnetic moment M of the atom, explained in the last section. We have remarked that for the Zeeman-effect and the Stark-effect, when we bring an external magnetic or electric field in a definite direction in space, we have to introduce a new quantum number connected with the azimuthal angle ϕ of the electron round the axis of the field. Similarly, the magnetic field due to the atomic magnetism, introduces the quantum number to which the multiplet structure is due. This number 'm' is called the "magnetic quantum number".

So far, we see that the negative energy W of a stationary state depends on three quantum numbers n, l, m. But in the course of investigations on the multiplet structure of series lines, it was found that we need yet one more, a fourth, quantum number to explain them.

Thus consider the effect of an external magnetic field on an atom. In the normal Zeeman-effect, it is found that each line of the hydrogen spectrum is resolved into three components under the influence of the external field. But the case is different for the alkali atoms. If the magnetic field is weak, each line of the multiplet is resolved into more than three lines. The displacement of these new components is proportional to H, but different for the various members of the multiplet. This phenomenon is called the "anomalous Zeeman-effect".

It is found further that if the magnetic field is gradually increased, the resolution is gradually changed; and finally when the field becomes strong, there remain only three components of each line of the multiplet, as in the normal Zeeman-effect. This change on passing from a weak to a strong field, is called the "Paschen-Back" effect.

The fourth quantum number mentioned above, was introduced to explain these phenomena. It is denoted by s, and is supposed to have only the two values $+\frac{1}{2}$ and $-\frac{1}{2}$. It was at first assumed by Landé and others that this fourth quantum number s was connected with the angular momentum of the core electrons, but in 1924 Pauli showed that this cannot be the case. He pointed out that if s were connected with the core electrons, then the nature of the Zeeman-effect would depend upon the atomic number N of the atom. But experiment shows that there is no such dependence of the Zeeman-effect upon N. Pauli concluded, therefore, that the quantum number s is associated with the series electron, just as the other three numbers are associated with it.

This view was further strengthened by the hypothesis of the "spinning electron," proposed by two Dutch Physicists, G. E. Uhlenbeck and S. Goudsmit, in October 1925. They suggested that the electron should not be considered to be a point charge, as hitherto done in Bohr's theory, but must be taken to have a finite extension. Further, it should be assumed to be spinning about an axis fixed in itself, while at the same time moving in its orbit round the nucleus. The angular momentum of the electron is taken to be $s\frac{h}{2\pi}$, where $s = +\frac{1}{2}$ or $-\frac{1}{2}$, according as the spin is in one direction or the other. The hypothesis of the spin of the electron coupled with the relativity theory gave a satisfactory explanation of the spin doublets. But in 1928, Dirac showed that a systematic relativity theory of the quantum phenomena accounts for the spin effect quite naturally, without the arbitrary model of the spinning electron (Chap. XI).

5·8 (3). *Pauli's Exclusion Principle.*

We have seen in the last section that each stationary state of an atom is characterised by the following four quantum numbers:

(1) The principal quantum number n, which can have any of the integral values 1, 2, 3, $\cdots$, and which in a transition can change by any amount.

(2) The azimuthal quantum number l, which, for a fixed n, can have only the following n values (25) § 5·7:

$$0, 1, 2, \cdots, (n-1),$$

and which, according to (30 *a*) § 5·7, can change in a transition by -1 or $+1$ only.

(3) The magnetic quantum number m, which, according to (1) § 5·8, has, for given values of n and l, the following $(2l+1)$ values:

$$-l, -(l-1), \cdots, -1, 0, 1, \cdots, (l-1), l,$$

and which can change in a transition by -1, 0, or $+1$ only.

(4) The spin quantum numbers s, which can have only the two values $-\frac{1}{2}$ and $+\frac{1}{2}$.

Thus the negative energy W, associated with a stationary state, is a function of the four numbers n, l, m and s, so that we can write

$$W = W(n, l, m, s).$$

Now, from the time of the Greek philosophers, it has been taken for granted that two different pieces of matter cannot occupy the same part of space at the same time. Generalising this idea, W. Pauli put forward, in 1925, as a fundamental postulate of the quantum theory, that two electrons not only cannot be at the same point simultaneously, they cannot even be in the same orbit. As each orbit is characterised by the four quantum numbers explained above, this means simply that no two electrons can have the same four quantum numbers. In other words, each stationary state of the quantum theory is defined completely and uniquely by a given set of values of the four numbers n, l, m, s. This is Pauli's "Exclusion Principle". It is also called Pauli's "Equivalence Principle" or "Pauli-Verbot".

This is a most useful principle, and can be applied to find out the number and arrangement of the extra-nuclear electrons in the atoms of the various chemical elements.

5·9. *Bohr's Theory of the Structure of Atoms, and the Periodic Table of the Elements.*

In § 3·1 we said that the chemical elements can be arranged according to their increasing atomic weight, beginning with the lightest of them, *viz.*, hydrogen, and ending with the heaviest, *viz.*, uranium. In the middle of the nineteenth century the Russian chemist, Mendéleef, suggested that as regards their chemical and physical properties, such as valency and the emission of spectral lines, these chemical elements, which were supposed to be 92 in number, can be divided into seven periods, having more or less resemblance to each other, and containing 2, 8, 8, 18, 18, 32, and 6 elements respectively. The properties in each period are repeated, and each increase in the length of the period brings with it a number of elements with properties different from those of the preceding periods.

When Bohr first proposed his quantum theory of the atom in 1913, he suggested that the chemical and physical properties depended on the peripheral electrons. Since that time it was believed that the repetition of the chemical and physical properties with these periods, and the increasing length of the period, is due to a regular building up of the atoms. But it was only in 1922 that Bohr was able to give a theory of the periodic table. His considerations were based on a detailed knowledge of the optical and X-ray spectra, the Zeeman-effect, and other phenomena. The theory was given a firm basis only after the publication of Pauli's Exclusion Principle in 1925.

In the last section we saw that the energy associated with any stationary state is characterised by four quantum numbers n, l, m and s. We know also that only those configurations are most probable for which the energy is a minimum. Thus at each state the electrons in the neutral atom would be in the orbits of the lowest possible energy compatible with the condition of stability. These ideas together with the exclusion principle, provide the necessary clue for the fixing up of the arrangement of electrons in the various atoms.

When the theory is worked out, it is found that in every atom there are two orbits in which the energy is equal and lower than in any other orbit. These are nearest to the nucleus, and their group is spoken of as the K-shell or K-ring.

Next to these come eight orbits of equal energy, but substantially higher than that of the K-shell. This group is called the L-shell. After the L-shell we get 18 orbits of equal but still higher energy, and the group is known as the M-shell. Similarly, the N-shell contains 32 orbits, and the O-shell 50 orbits. The atoms are successively built up by the addition of one peripheral electron to those of the previous atom. The electrons go into the various orbits beginning with the K-shell, and when one shell is complete, the additional electron of the next higher atom must go into an orbit belonging to the next higher shell.

From Bohr's theory of the hydrogen atom, we know that the value of the energy in a stationary state depends mainly on the principal quantum number n, and it is only the slight differences in energy due to the finer features which bring about the introduction of the other quantum numbers.

Thus we see that the various shells are characterised by the values of the principal quantum number $n = 1, 2, 3, 4, 5$, etc., giving the K – , L – , M – , N – , O – , P – , Q – shells respectively.

From (25) § 5·7 we know that for a fixed value of n, the azimuthal quantum number has the n values:

$$0, 1, 2, \cdots, (n - 1).$$

Therefore each shell has n sub-shells characterised by the value of l. These sub-shells are distinguished by suffixes. Thus the three sub-shells of the M-shell ($n = 3$), corresponding to the values 0, 1, 2 of l, are denoted by M_1, M_2, M_3 respectively.

Now we must consider the other two quantum numbers m and s. We know that for given values of n and l, the number m can have only the $(2l + 1)$ values:

$$-l, -(l - 1), \cdots, -1, 0, +1, \cdots, (l - 1), l, \tag{2}$$

and for each of these m, the fourth number s can assume the two values $+\frac{1}{2}$ and $-\frac{1}{2}$. For any sub-shell n_l the number of orbits would therefore be $2\ (2l + 1)$. And since each different orbit may contain an electron, the number of electrons in each sub-shell will be at the most $2\ (2l + 1)$. The number of electrons in the various sub-shells should be therefore

$$2, 6, 10, 14, \cdots. \tag{3}$$

The number of electrons in a shell is then

$$\sum_{l=0}^{n-1} 2\ (2l + 1) = 2n^2. \tag{4}$$

This gives the maximum possible number of electrons in each sub-shell and shell, calculated from Pauli's Exclusion Principle. But the electron configuration must satisfy also the second condition that the energy of the atom shall be a minimum. When this second condition is taken into account, it is found that the higher shells do not have the maximum number of electrons calculated from the above theory, but that one shell begins to form before the previous shell is complete.

The prevalent grouping of the electrons into shells and sub-shells for the more prominent chemical elements is given in the following table. We wish to remark here that 92 is not the upper limit for the number of elements. In fact E. Fermi recently created in his laboratory the element No. 93, which is, however, extremely unstable. Elements No. 94 and 95 have also been reported.

TABLE I.

The Electronic Structure of Atoms.

	$n=1$ K	2 L		3 M			4 N				5 O			6 P		7 Q
	$l=0$	0	1	0	1	2	0	1	2	3	0	1	2	0	1	0
1. H	1															
2. He	2															
6. C	2	2	2													
7. N	2	2	3													
8. O	2	2	4													
9. F	2	2	5													
10. Ne	2	2	6													
11. Na	2	2	6	1												
12. Mg	2	2	6	2												
16. S	2	2	6	2	4											
17. Cl	2	2	6	2	5											
18. A	2	2	6	2	6											
25. Mn	2	2	6	2	6	5	2									
26. Fe	2	2	6	2	6	6	2									
27. Co	2	2	6	2	6	7	2									
28. Ni	2	2	6	2	6	8	2									
36. Kr	2	2	6	2	6	10	2	6								
46. Pd	2	2	6	2	6	10	2	6	10							
47. Ag	2	2	6	2	6	10	2	6	10		1					
54. Xe	2	2	6	2	6	10	2	6	10		2	6				
86. Rn	2	2	6	2	6	10	2	6	10	14	2	6	10	2	6	
88. Ra	2	2	6	2	6	10	2	6	10	14	2	6	10	2	6	2

CHAPTER VI.

HEISENBERG'S QUANTUM MECHANICS.

6·1. *Critique of Bohr's Theory.*

IN the last chapter we gave an account of the origin and development of Bohr's quantum theory, and pointed out its culmination in the explanation of the periodic table of chemical elements published by Bohr in 1922. It became increasingly apparent, however, that this theory was not quite complete. It could explain the Zeeman-effect of the hydrogen atom, but gave wrong results for the anomalous Zeeman-effect. Further, the energy levels of the hydrogen molecule ion, of the normal and excited states of the helium atom, of the normal states of the hydrogen molecule, were calculated by many people during the years 1923–25, but it was found that the results did not at all agree with experiments. The quantum theory broke down for the problem of the crossed fields, *i.e.*, the problem of calculating the spectrum of an atom under simultaneous electric and magnetic fields in different directions. The phenomena of dispersion, *i.e.*, of the collision of light with matter, and also those of the collision of electrons with gas atoms could not be accounted for successfully on Bohr's theory.

Indeed, Bohr himself was never in doubt about the incompleteness of his theory of atom models and electron orbits. Max Born, W. Pauli, and many other workers of great repute were giving voice to their belief that a reformulation of the whole theory was urgently needed. This was felt to be the more necessary because the quantum theory was not a coherent whole; there was an unsatisfactory looseness about its principles. The quantum conditions, the frequency relation and the correspondence principle are not integral parts of the theory, but are added to it as an after-thought, as it were. The theory lacked the closeness and unity of the classical mechanics, which did not require any external ideas for the explanation of the phenomena with which it dealt. Several workers were seeking for a rational and unified formulation of the quantum laws, but it was reserved for a very young German physicist, W. Heisenberg, to provide the key to the solution.

6· 2. *Heisenberg's Theory.*

Heisenberg pointed out in 1925, that any rational formulation of the quantum mechanics should concern itself directly with the experimentally observable quantities such as the frequencies and intensities of the spectrum. The position of the electron in the orbit, the period of revolution and the amplitudes can never be expected to be experimentally observed, and should not be used directly in the mechanics.

In the course of the last chapter, we have seen that the fundamental ideas of Bohr's theory are:

(1) the discrete set of stationary states characterised by the values E_n $(n = 1, 2, \cdots)$, of the energy;

(2) the frequency ν of the energy radiation emitted during the transition from a state n to another state n', given by Bohr's frequency relation $\nu = \frac{E_n - E_{n'}}{h}$;

(3) the orbits of the electrons, selected by the quantum conditions $\oint p_r \, dq_r = n_r h, \quad (r = 1, 2, \cdots, u)$;

(4) the assumption that the classical laws hold for the description of motion in the stationary states, although they fail completely during transitions;

(5) the correspondence principle, that the classical theory gives the right results in the limiting case of large quantum numbers.

The existence of the discrete set of stationary states for the atom has been confirmed more or less directly by the experiments of J. Franck and others on collision phenomena. The frequency relation can be taken to be an immediate consequence of the Ritz combination principle, which has been abundantly demonstrated during more than half a century. These two conceptions, therefore, must find a place in the new mechanics.

The case of the electron orbits is, however, quite different. It is almost certain that they can never be observed experimentally. Accordingly, the orbits should not enter directly in the new theory. Heisenberg thus gave up all mechanical models of the atom, just as previously, Einstein had given up the omnipresent ether of the classical electrodynamics. Moreover, the quantum conditions, which provided the selection rules for these electron orbits, were completely artificial; they could not be interpreted rationally in terms of the ordinary corpuscular conceptions of the nature of matter. Their only justification was their agreement with the facts of experiment in some cases. All that the adiabatic principle of Ehrenfest and the correspondence principle of Bohr could do for the quantum conditions, was to make them appear plausible to a certain extent.

As regards the classical mechanics, Heisenberg put forward that it is not the equations of classical mechanics that are at fault, but that the mathematical operations by which physical results are deduced from these equations require modification.

Heisenberg considered a simply periodic dynamical system, with the orbital frequency $\omega = 2\pi\nu$. The co-ordinate q can then be represented by a Fourier series.

$$\begin{aligned} q(t) &= a_0 + a_1 \cos 2\pi\nu t + a_2 \cos 2(2\pi\nu t) + \cdots \\ &\quad + b_1 \sin 2\pi\nu t + b_2 \sin 2(2\pi\nu t) + \cdots \\ &= a_0 + \tfrac{1}{2}(a_1 - ib_1)\, e^{2\pi i\nu t} + \tfrac{1}{2}(a_2 - ib_2)\, e^{2(2\pi i\nu t)} + \cdots \\ &\quad + \tfrac{1}{2}(a_1 + ib_1)\, e^{-2\pi i\nu t} + \tfrac{1}{2}(a_2 + ib_2)\, e^{-2(2\pi i\nu t)} + \cdots \\ &= q_0 + q_1 e^{2\pi i\nu t} + q_2 e^{2(2\pi i\nu t)} + q_3 e^{3(2\pi i\nu t)} + \cdots \\ &\quad + q_{-1} e^{-2\pi i\nu t} + q_{-2} e^{-2(2\pi i\nu t)} + q_{-3} e^{-3(2\pi i\nu t)} + \cdots \\ &= \sum_{\alpha=-\infty}^{+\infty} q_\alpha\, e^{\alpha(2\pi i\nu t)}, \end{aligned} \tag{1}$$

where

$$q_\alpha = \tfrac{1}{2}(a_1 - i\, b_1), \quad \text{(for } \alpha \text{ positive integral)},$$

and

$$q_{-\alpha} = \tfrac{1}{2}(a_1 + i\, b_1).$$

If we denote the complex conjugate of q by q^*, then we see from the definitions of q_α and $q_{-\alpha}$ that

$$q_{-\alpha} = q_\alpha^*. \tag{2}$$

Heisenberg assumed that in the new quantum mechanics, the coordinate q can be represented by harmonic components of the form $qe^{2\pi i\nu t}$. Since in the quantum theory the observable frequencies and intensities of the spectral lines depend on a transition from a state n to another state m, q and ν should be characterised by the two integers n and m instead of the one integer α of the classical equation (1). Now, according to the

Ritz combination principle, the frequency of a spectral line is given as a difference of two terms of the set:

$$\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3, \cdots \mathrm{T}_n, \cdots,$$

so that

$$\nu_{nm} = \mathrm{T}_n - \mathrm{T}_m \tag{3}$$

is the frequency of the transition from the state n to m.

Thus

$$\begin{aligned}\nu_{nm} + \nu_{mk} &= (\mathrm{T}_n - \mathrm{T}_m) + (\mathrm{T}_m - \mathrm{T}_k)\\ &= \mathrm{T}_n - \mathrm{T}_k = \nu_{nk},\end{aligned}$$

which corresponds to the obvious classical relation

$$\alpha\nu + \beta\nu = (\alpha + \beta)\,\nu.$$

Further, we have

$$\nu_{nn} = \mathrm{T}_n - \mathrm{T}_n = 0, \tag{4}$$

which embodies the fact that no radiation is emitted so long as the system remains in the same state. Similarly

$$\begin{aligned}\nu_{mn} &= \mathrm{T}_m - \mathrm{T}_n = -\,(\mathrm{T}_n - \mathrm{T}_m)\\ &= -\,\nu_{nm},\end{aligned} \tag{5}$$

showing that the frequency of the radiation emitted during the transition $n \rightarrow m$, is the same as the frequency of the radiation absorbed during the reverse transition $m \rightarrow n$. It must be remarked that negative frequencies signify absorption in the same way as positive frequencies signify emission.

Heisenberg found, therefore, that instead of the Fourier series (1), the co-ordinate q should be represented in the new mechanics corresponding to all possible transitions n (= 1, 2, 3) to m (= 1, 2, 3, $\cdots$), by the following two dimensional table:

$$q = \begin{pmatrix} q_{11}\,e^{2\pi i\nu_{11}t}, & q_{12}\,e^{2\pi i\nu_{12}t}, & q_{13}\,e^{2\pi i\nu_{13}t}, & \cdots\\ q_{21}\,e^{2\pi i\nu_{21}t}, & q_{22}\,e^{2\pi i\nu_{22}t}, & q_{23}\,e^{2\pi i\nu_{23}t}, & \cdots\\ q_{31}\,e^{2\pi i\nu_{31}t}, & q_{32}\,e^{2\pi i\nu_{32}t}, & q_{33}\,e^{2\pi i\nu_{33}t}, & \cdots\end{pmatrix}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

or more, shortly by

$$q = (q_{nm}\,e^{2\pi i\nu_{nm}t}). \tag{6}$$

But from (4) we know that $\nu_{11} = \nu_{22} = \nu_{33} \cdots = 0$, so that all the diagonal terms are constants $q_{11}, q_{22}, \cdots$. Further, since the dynamical variable q is real, and since $\nu_{mn} = -\,\nu_{nm}$, we must demand that q_{mn} should be the complex conjugate of q_{nm}, so that

$$q_{mn} = q^{*}_{nm}, \tag{7}$$

and

$$q_{nm}\, q_{mn} = |q_{nm}|^2. \tag{8}$$

In the classical theory, if two variables x and y belong to the same dynamical system with the orbital frequency $\omega = 2\pi\nu$, then

$$x = \sum_{\alpha=-\infty}^{+\infty} x_\alpha\, e^{\alpha(2\pi i\nu t)},$$

$$y = \sum_{\beta=-\infty}^{+\infty} y_\beta\, e^{\beta(2\pi i\nu t)},$$

and

$$\begin{aligned} xy &= \{\sum_\alpha x_\alpha\, e^{\alpha\,(2\pi i\nu t)}\}\{\sum_\beta y_\beta\, e^{\beta\,(2\pi i\nu t)}\}, \\ &= \sum_\alpha \sum_\beta x_\alpha y_\beta\ e^{2\pi i\,(\alpha+\beta)\,\nu t}. \end{aligned}$$

Now let

$$\alpha + \beta = \gamma, \qquad \beta = \gamma - \alpha\,;$$

then

$$\begin{aligned} xy &= \sum_\gamma \{\sum_\alpha x_\alpha y_{\gamma-\alpha}\}\, e^{2\pi i\gamma\nu t}, \\ &= \sum_\gamma\, (xy)_\gamma\, e^{2\pi i\gamma\nu t}, \end{aligned} \tag{9}$$

where

$$(xy)_\gamma = \sum_\alpha x_\alpha y_{\gamma-\alpha}. \tag{10}$$

We see therefore that the same frequencies ν, 2ν, 3ν, $\cdots$ enter into the product xy as into x and y alone.

Heisenberg wanted to find a rule of multiplication for the quantum dynamical variables so that no new frequencies should appear in the product. Thus if

$$x = (x_{nm}\, e^{2\pi i\nu_{nm}t}),$$

$$y = (y_{nm}\, e^{2\pi i\nu_{nm}t}),$$

then Heisenberg found that it is only when we stipulate that the $n\,m$ component of xy should be calculated by the rule :

$$(xy)_{nm} = \sum_k x_{nk}\, e^{2\pi i\nu_{nk}t}\, y_{km}\, e^{2\pi i\nu_{km}t}, \tag{11}$$

that no new freqencies appear in the product. In fact, since on account of (3),

$$\begin{aligned} \nu_{nk} + \nu_{km} &= (\mathrm{T}_n - \mathrm{T}_k) + (\mathrm{T}_k - \mathrm{T}_m), \\ &= (\mathrm{T}_n - \mathrm{T}_m) = \nu_{nm}, \end{aligned}$$

we have from (11)

$$(xy)_{nm} = \sum_k x_{nk} y_{km} \, e^{2\pi i(\nu_{nk}+\nu_{km})t}$$

$$= \sum_k x_{nk} y_{km} \, e^{2\pi i \nu_{nm} t}. \tag{12}$$

According to this rule, the nm component of the product yx would be given by

$$(yx)_{nm} = \sum_k y_{nk} x_{km} \, e^{2\pi i \nu_{nm} t}. \tag{13}$$

In the general case, the sums $\sum_k x_{nk}\, y_{km}$ and $\sum_k y_{nk}\, x_{km}$ will evidently not be equal, and therefore $(xy)_{nm} \neq (yx)_{nm}$ for all values of n and m. Thus the corresponding components in the tables for xy and yx are not equal, and therefore

$$xy \neq yx. \tag{14}$$

This fact, that the commutative law of multiplication does not hold in the new quantum mechanics, was mentioned by Heisenberg explicitly in his very first paper published in July 1925. Max Born and P. Jordan pointed out in September 1925, that the Heisenberg multiplication rule (11) is exactly the same as that for matrices, well known to the mathematicians in algebraic theory. Thus, the dynamical variables in quantum mechanics are " matrices ".

6·3. *The Algebra of Matrices.*

Suppose that x_1, x_2 are two variables, which are considered to be components of a two dimensional vector x. In the same way we shall say that y_1, y_2 are components of a vector y ; z_1, z_2 are components of a vector z ; etc.

We make the following transformation of variables:

$$y_1 = a_{11} x_1 + a_{12} x_2, \; y_2 = a_{21} x_1 + a_{22} x_2. \tag{1}$$

We arrange the coefficients of this transformation in a square table of two rows and two columns :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{2}$$

This table is called a " matrix " of the 2nd order, and is usually denoted by a. Then the two equations in (1) can be written symbolically as a sort of vector equation

$$y = ax. \tag{3}$$

We transform the variables again from y_1, y_2 to z_1, z_2 by means of the transformation

$$z_1 = b_{11}y_1 + b_{12}y_2, \; z_2 = b_{21}y_2 + b_{22}y_2, \tag{4}$$

or, shortly by

$$z = by, \tag{5}$$

where b is the transformation coefficients matrix

$$b = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}. \tag{6}$$

Substituting the values of y_1, y_2 from (1) in (4), we can find the new variables z_1, z_2 in terms of the original variables x_1, x_2 as

$$z_1 = c_{11}x_1 + c_{12}x_2, \quad z_2 = c_{21}x_1 + c_{22}x_2, \tag{7}$$

where it is easily calculated that

$$c_{11} = b_{11}a_{11} + b_{12}a_{21}, \; c_{12} = b_{11}a_{12} + b_{12}a_{22},$$

$$c_{21} = b_{21}a_{11} + b_{22}a_{21}, \; c_{22} = b_{21}a_{12} + b_{22}a_{22}. \tag{8}$$

We can write (7) shortly in the symbolical form

$$z = cx, \; c = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}. \tag{9}$$

Substituting (3) in (5) we have

$$z = by = b\,\{ax\} = bax, \tag{10}$$

which means that first the transformation a is applied on x, and then the transformation b is applied on ax. Comparing (9) and (10), we get

$$c = ba, \tag{11}$$

that is to say that the product of two matrices is a third matrix whose elements are calculated according to the formulæ (8). That the product ab is not the same as ba can be readily seen by applying the transformations in the other order. Thus let

$$y_1' = b_{11}x_1 + b_{12}x_2, \quad y_2' = b_{21}x_1 + b_{22}x_2, \tag{12}$$

or, symbolically,

$$y' = bx. \tag{13}$$

Applying now the transformation a, we get

$$z_1' = a_{11}y_1' + a_{12}y_2', \quad z_2' = a_{21}y'_1 + a_{22}y_2', \tag{14}$$

or

$$z' = ay. \tag{15}$$

Substituting (12) in (14), we get

$$z_1' = d_{11}x_1 + d_{12}x_2, \quad z_2' = d_{21}x_1 + d_{22}x_2, \tag{16}$$

where

$$d_{11} = a_{11}b_{11} + a_{12}b_{21}, \; d_{12} = a_{11}b_{12} + a_{12}b_{22},$$

$$d_{21} = a_{21}b_{11} + a_{22}b_{21}, \; d_{22} = a_{21}b_{12} + a_{22}b_{22}. \quad (17)$$

Thus

$$z' = dx, \; d = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}. \quad (18)$$

Substituting (13) in (15), we get

$$z' = ay' = a\,\{bx\} = abx. \quad (19)$$

From (18) and (19) we have

$$d = ab. \quad (20)$$

Comparing (8) and (17), we see that in general all the four c's, *viz.*, c_{11}, c_{12}, c_{21}, c_{22}, are not the same as all the four d's, *viz.*, d_{11}, d_{12}, d_{21}, d_{22} respectively, so that $c \neq d$, *i.e.*,

$$ba \neq ab. \quad (21)$$

Thus we see that the commutative law of multiplication does not hold for matrices in general.

Now, let A be the determinant formed of the coefficient of the matrix A, and suppose that

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0, \quad (22)$$

then from (1), solving for x_1, x_2 in terms of y_1, y_2, we get

$$x_1 = \frac{a_{22}}{A}\,y_1 - \frac{a_{12}}{A}\,y_2, \quad x_2 = \frac{a_{21}}{A}\,y_1 - \frac{a_{11}}{A}\,y_2. \quad (23)$$

This is called the inverse transformation to (1), and the matrix of the coefficients in (23) is called the inverse matrix to a and is denoted by a^{-1}, so that we can write (23) symbolically:

$$x = a^{-1}y, \; a^{-1} = \begin{pmatrix} \frac{a_{22}}{A}, & -\frac{a_{12}}{A} \\ \frac{a_{21}}{A}, & -\frac{a_{11}}{A} \end{pmatrix}. \quad (24)$$

It can be verified by actually carrying out the double transformations, that

$$a^{-1}\,ax = a\,a^{-1}\,x = x,$$

where

$$a^{-1}a = aa^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (25)$$

The matrix (25) is called the unit matrix of the 2nd order, and we see that every matrix "commutes" with its inverse, in other words, that the commutative law of multiplication holds when a matrix is multiplied by its inverse.

All these results can be generalised to matrices of any order. Thus

$$a = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{pmatrix} = (a_{mn}) \tag{26}$$

is called a finite matrix of order s. Each of the a_{mn} is called an element of the matrix. The elements $a_{11}, a_{22}, \cdots, a_{ss}$ are said to constitute the leading diagonal. If $s = \infty$, the matrix is said to be infinite.

A matrix a is said to be equal to another matrix b of the same order, if $a_{mn} = b_{mn}$ for all m and all n.

The matrix c is said to be the sum of two matrices a, b of the same order, if

$$c_{mn} = a_{mn} + b_{mn} \begin{pmatrix} m = 1, 2, \cdots, s \\ n = 1, 2, \cdots, s \end{pmatrix};$$

this is expressed by the equation $c = a + b$.

The matrix c is said to be the product of two matrices a, b of the same order, if

$$c_{mn} = \sum_{k=1}^{s} a_{mk} b_{kn}, \tag{27}$$

and this is expressed by the equation $c = ab$. This rule shows evidently that

$$ba \neq ab,$$

i.e., the commutative law does not hold in general. If in a particular instance the two products ba and ab are the same, the two matrices a, b are said "to be commutable," or simply "to commute".

It is easy to see that

$$(ab)\, c = a\, (bc),$$

$$a\, (b + c) = ab + ac,$$

i.e., the associative and distributive laws hold in general.

It should always be remembered that the sum, difference, product of two or any number of matrices, will always be a matrix, and not an ordinary number. The matrix whose each and every element is zero, *i.e.*, for which $a_{mn} = 0$ for all m and all n, is called a "zero matrix," and is denoted simply by O.

It is easy to give examples of two matrices a, b of which each one is different from a zero matrix, and yet the product ab is a zero matrix. This is simply a consequence of the definitions and the rule of multiplication, and should not cause any surprise to the reader. We define a set of quantities δ_{mn} as

$$\begin{aligned} \delta_{mn} &= 1 \text{ for } n = m, \\ &= 0 \text{ for } n \neq m. \end{aligned} \tag{28}$$

The matrix whose elements are δ_{mn}, *i.e.*, the matrix in which all the elements in the leading diagonal are 1, and all the other elements are zero, is called a "unit matrix". We shall denote it by δ, so that

$$\delta = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} = (\delta_{mn}). \tag{29}$$

Now let $c = a\delta$, where a is any matrix and δ is unit matrix of the same order. Then from (27)

$$c_{mn} = \sum_k a_{mk}\, \delta_{kn} = a_{mn}.$$

Thus $c = a$. Similarly, it can be shown that if $b = \delta a$, then $b = a$. Hence we get the result

$$a\delta = \delta a = a. \tag{30}$$

Incidentally, we see from (30) that δ commutes with every matrix of the same order.

The matrix $E_{mn}\,\delta_{mn}$, *i.e.*, one which has non-vanishing elements only in the leading diagonal, so that all the elements not in the leading diagonal are zero, is called a "diagonal matrix", or a "scalar matrix". Such a matrix is of the type:

$$E = \begin{pmatrix} E_{11} & 0 & 0 & 0 & \cdots \\ 0 & E_{22} & 0 & 0 & \cdots \\ 0 & 0 & E_{33} & 0 & 0 \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = (E_{mn}\, \delta_{mn}). \tag{31}$$

It is easy to verify that a diagonal matrix commutes with every other matrix, so that $Ea = aE$.

The elements in the leading diagonal of a diagonal matrix are called the "Eigenvalues" of this matrix.

The inverse matrix a^{-1} of a is defined by the relation

$$a^{-1}a = \delta. \tag{32}$$

Then we have

$$a\,(a^{-1}a) = a\delta = a = \delta a,$$

so that on account of the associative property.

$$(a\,a^{-1})\,a = \delta a.$$

Consequently

$$a\,a^{-1} = \delta = a^{-1}a. \tag{33}$$

Equation (33) shows, as already mentioned in (25) above, that the inverse of every matrix commutes with the matrix itself.

On account of the non-commutative character of multiplication, the division of a matrix b by a matrix a is not unique. There would be in general two different quotients x and y, given by the two equations

$$\left.\begin{aligned} ax &= b, \\ ya &= b. \end{aligned}\right\} \tag{34}$$

If we multiply the first equation by the inverse matrix a^{-1} on the left, we get

$$a^{-1}ax = a^{-1}b, \text{ or } \delta x = a^{-1}b, \text{ or } x = a^{-1}b.$$

Similarly, if we multiply the second equation in (34) by a^{-1} on the right, we get

$$ya\,a^{-1} = ba^{-1}, \text{ or } y\delta = ba^{-1}, \text{ or } y = ba^{-1}.$$

Thus the two quotients are $a^{-1}b$ and ba^{-1}, which are generally different.

If $a_{mn} = a_{nm}$ for all m and all n, the matrix a is called a "symmetrical" matrix; but if $a_{mn} = -a_{nm}$, a is called a "skew-symmetrical" matrix.

Let a^*_{mn} denote the quantity conjugate to a_{mn}. Now, if $a_{nm} = a^*_{mn}$, the matrix a is called a 'Hermitian' matrix. The properties of such matrices were extensively studied by the French mathematician Hermite. They are most important for us, because on account of (7) § 6· 2, we find that the quantum mechanical variables must be matrices of the Hermitian type, if they are to be real. Thus

$$a_{nm} = a^*_{mn}. \tag{35}$$

Having discovered that the quantum variables are matrices Heisenberg, Born and Jordan, had to resort to a sort of differentiation of matrices, in order to be able to develope their mechanics. But this was rather a lengthy and cumbersome method. A few days later, in November 1925, P. A. M. Dirac of Cambridge, found that if the "Poisson Bracket" expressions of the classical mechanics are introduced, we can avoid having to define a differentiation of the matrices.

6·4. *Poisson Brackets.*

Consider a dynamical system having s degrees of freedom. The generalised co-ordinates are $q_1, \cdots, q_s$, and the conjugate momenta $p_1, \cdots, p_s$. In the classical mechanics, any two variables x, y are said to have a "Poisson Bracket", given by:

$$[x, y] = \sum_{r=1}^{s} \left(\frac{\partial x}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial y}{\partial q_r}\right). \tag{1}$$

The importance of a Poisson Bracket* expression lies in the fact that they remain invariant under a canonical transformation. We shall prove this theorem for the case of a single degree of freedom. The argument can be immediately extended to several degrees.

We saw in § 2· 5 that if we change from the variables p, q to the new variables, P, Q, given by $P = P(p, q)$, $Q = Q(p, q)$, then the new variables will be canonical if

$$p\,dq + Q\,dP = dW, \tag{2}$$

where W is the transformation function, and where we have assumed that the Hamiltonian does not contain the time explicitly, so that $H(p, q) = K(P, Q)$.

Equation (2) can be written in the expanded form:

$$\left(p + Q\frac{\partial P}{\partial q}\right) dq + Q\frac{\partial P}{\partial p}\,dp = \frac{\partial W}{\partial q}\,dq + \frac{\partial W}{\partial p}\,dp, \tag{3}$$

which gives

$$\frac{\partial W}{\partial q} = p + Q\frac{\partial P}{\partial q}, \quad \frac{\partial W}{\partial p} = Q\frac{\partial P}{\partial p}. \tag{4}$$

Since we must have $\dfrac{\partial^2 W}{\partial p \partial q} = \dfrac{\partial^2 W}{\partial q \partial p}$, we get from (4):

$$\frac{\partial}{\partial p}\left(p + Q\frac{\partial P}{\partial q}\right) = \frac{\partial}{\partial q}\left(Q\frac{\partial P}{\partial p}\right),$$

---

* Sometimes we shall write simply P.B. for Poisson Bracket.

or

$$1 + \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} + Q\frac{\partial^2 P}{\partial p \partial q} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} + Q\frac{\partial^2 P}{\partial q \partial p},$$

so that

$$[Q, P] = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} = 1. \tag{5}$$

This shows that if the transformation to the new variables is canonical, the Jacobian (functional determinant) must be equal to unity.

Now let x, y be any two functions of p, q, and consequently of P,Q. Then the P.B. of x, y is

$$[x, y] = \begin{vmatrix} \frac{\partial x}{\partial q} & \frac{\partial y}{\partial q} \\ \frac{\partial x}{\partial p} & \frac{\partial y}{\partial p} \end{vmatrix} = \begin{vmatrix} \frac{\partial Q}{\partial q} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial P}{\partial p} \end{vmatrix} \begin{vmatrix} \frac{\partial x}{\partial Q} & \frac{\partial y}{\partial Q} \\ \frac{\partial x}{\partial P} & \frac{\partial y}{\partial P} \end{vmatrix}$$

$$= \frac{\partial x}{\partial Q}\frac{\partial y}{\partial P} - \frac{\partial x}{\partial P}\frac{\partial y}{\partial Q}, \tag{6}$$

on account of (5). This proves that the P.B. expression is invariant for a canonical transformation.

Reverting to the general case of several degrees of freedom, if x is any function of the q_r's and p_r's, we have :

$$\frac{\partial x}{\partial q_r} = \sum_{k=1}^{s} \left(\frac{\partial x}{\partial q_k}\frac{\partial p_r}{\partial p_k} - \frac{\partial x}{\partial p_k}\frac{\partial p_r}{\partial q_k}\right) = [x, p_r],$$

$$(r = 1, 2, \cdots, s) \tag{7}$$

$$\frac{\partial x}{\partial p_r} = -\sum_{k=1}^{s} \left(\frac{\partial x}{\partial q_k}\frac{\partial q_r}{\partial p_k} - \frac{\partial x}{\partial p_k}\frac{\partial q_r}{\partial q_k}\right) = -[x, q_r].$$

Thus we find that differentiation *w. r.* to the variables q_r, p_r can be replaced by P.B. expressions. This makes it possible to write Hamilton's canonical equations of motion in the form :

$$\dot{q}_r = \frac{\partial H}{\partial p_r} = -[H, q_r], \qquad (r = 1, 2, \cdots, s). \tag{8}$$

$$\dot{p}_r = -\frac{\partial H}{\partial q_r} = -[H, p_r]. \tag{9}$$

This form of the equations of motion verifies immediately that they are invariant for a canonical transformation.

We can immediately prove from definition the following properties of the Poisson Brackets.

(*a*) $$[y, x] = \sum_{r=1}^{s} \left(\frac{\partial y}{\partial q_r}\frac{\partial x}{\partial p_r} - \frac{\partial y}{\partial p_r}\frac{\partial x}{\partial q_r}\right),$$
$$= -\sum_{r=1}^{s} \left(\frac{\partial x}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial y}{\partial q_r}\right),$$
$$= -[x, y].$$

(*b*) If c is an ordinary number, then $\frac{\partial c}{\partial q_r} = 0, \frac{\partial c}{\partial p_r} = 0,$ therefore

$$[x, c] = \sum_{r=1}^{s} \left(\frac{\partial x}{\partial q_r}\frac{\partial c}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial c}{\partial q_r}\right) = 0.$$

(*c*) $$[q_m, q_n] = \sum_{r=1}^{s} \left(\frac{\partial q_m}{\partial q_r}\frac{\partial q_n}{\partial p_r} - \frac{\partial q_m}{\partial p_r}\frac{\partial q_n}{\partial q_r}\right) = 0$$

for all m and all n, because $\frac{\partial q_m}{\partial p_r} = \frac{\partial q_n}{\partial p_r} = 0.$

(*d*) $$[p_m, p_n] = \sum_{r=1}^{s} \left(\frac{\partial p_m}{\partial q_r}\frac{\partial p_n}{\partial p_r} - \frac{\partial p_m}{\partial p_r}\frac{\partial p_n}{\partial q_r}\right) = 0$$

for all m and all n, because $\frac{\partial p_m}{\partial q_r} = \frac{\partial p_n}{\partial q_r} = 0.$

(*e*) $$[q_m, p_n] = \sum_{r=1}^{s} \left(\frac{\partial q_m}{\partial q_r}\frac{\partial p_n}{\partial p_r} - \frac{\partial q_m}{\partial p_r}\frac{\partial p_n}{\partial q_r}\right)$$
$$= 0 \text{ if } m \neq n,$$
$$= 1 \text{ if } m = n.$$

This result can be expressed more systematically with the help of the δ-function defined in § 6· 4 (28) :

$$[q_m, p_n] = \delta_{mn}.$$

(*f*) $$[x_1 + x_2, y] = \sum_{r=1}^{s} \left\{\frac{\partial(x_1 + x_2)}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial(x_1 + x_2)}{\partial p_r}\frac{\partial y}{\partial q_r}\right\},$$
$$= \sum_{r=1}^{s} \left(\frac{\partial x_1}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial x_1}{\partial p_r}\frac{\partial y}{\partial q_r}\right) + \sum_{r=1}^{s} \left(\frac{\partial x_2}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial x_2}{\partial p_r}\frac{\partial y}{\partial q_r}\right),$$
$$= [x_1, y] + [x_2, y].$$

(*g*) Similarly, we get

$$[x, y_1 + y_2] = [x, y_1] + [x, y_2].$$

The properties (f) and (g) express the fact that the Poisson Bracket $[x, y]$ involves x and y linearly.

$$(h)\quad [x_1\, x_2, y] = \sum_{r=1}^{s} \left\{\frac{\partial(x_1\, x_2)}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial(x_1\, x_2)}{\partial p_r}\frac{\partial y}{\partial q_r}\right\},$$

$$= \left\{\sum_{r=1}^{s}\left(\frac{\partial x_1}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial x_1}{\partial p_r}\frac{\partial y}{\partial q_r}\right)\right\} x_2$$

$$+ x_1 \left\{\sum_{r=1}^{s}\left(\frac{\partial x_2}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial x_2}{\partial p_r}\frac{\partial y}{\partial q_r}\right)\right\},$$

$$= [x_1, y]\, x_2 + x_1\, [x_2, y].$$

$$(i)\quad [x, y_1\, y_2] = \sum_{r=1}^{s}\left\{\frac{\partial x}{\partial q_r}\frac{\partial(y_1\, y_2)}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial(y_1\, y_2)}{\partial q_r}\right\},$$

$$= \left\{\sum_{r=1}^{s}\left(\frac{\partial x}{\partial q_r}\frac{\partial y_1}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial y_1}{\partial q_r}\right)\right\} y_2$$

$$+ y_1 \left\{\sum_{r=1}^{s}\left(\frac{\partial x}{\partial q_r}\frac{\partial y_2}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial y_2}{\partial q_r}\right)\right\},$$

$$= [x, y_1]\, y_2 + y_1\, [x, y_2].$$

It is important in (h) and (i) to keep the factors in the order indicated here, because although in the classical theory it does not matter whether we write $[x_1, y].\, x_2$ or $x_2\, [x_1, y]$, but if we take them over into quantum mechanics, they would generally be different since x_1, x_2 and y are matrices.

If we keep this order, we can define the quantum Poisson Bracket so that it also has all these properties. They would be calculated formally as in the classical theory. Usually, the dynamical variables x, y are power series in the generalised co-ordinates q_r, p_r. Then we can build up $[x, y]$ from the values of (q_m, q_n), (p_m, p_n), (q_m, p_n) with the help of the above-mentioned properties. Thus

$$[q_1, p_1^2] = [q_1, p_1 p_1] = [q_1, p_1]\, p_1 + p_1\, [q_1, p_1].$$

But from (5) $[q_1, p_1] = 1$, therefore

$$[q_1, p_1^2] = 2\, p_1.$$

Similarly,

$$[q_1^2, p_1^2] = [q_1\, q_1, p_1^2] = [q_1, p_1^2]\, q_1 + q_1\, [q_1, p_1^2]$$
$$= 2\, p_1\, q_1 + 2 q_1 p_1.$$

$$[q_1^3, p_1^2] = [q_1^2\, q_1, p_1^2] = [p_1^2, p_1^2]\, q_1 + q_1^2\, [q_1, p_1^2]$$
$$= 2\,(p_1\, q_1 + q_1 p_1)\, q_1 + 2\, q_1^2\, p_1$$
$$= 2\, p_1\, q_1^2 + 2\, q_1 p_1 q_1 + 2\, q_1^2\, p_1,$$

and so on. Thus we see that the Poisson Bracket $[x, y]$ can be evaluated more conveniently by expressing it in terms of the simpler brackets.

We must give now a formal definition of the quantum Poisson Bracket. In order to do this, we shall calculate the value of the Poisson Bracket $[x_1x_2, y_1y_2]$ in two different ways.

$$\begin{aligned}[x_1x_2, y_1y_2] &= [x_1, y_1y_2]\,x_2 + x_1\,[x_2, y_1y_2],\\ &= \{[x_1, y_1]\,y_2 + y_1\,[x_1, y_2]\}\,x_2\\ &\quad + x_1\,\{[x_2, y_1]\,y_2 + y_1\,[x_2, y_2]\},\\ &= [x_1, y_1]\,y_2x_2 + y_1\,[x_1, y_2]\,x_2 + x_1\,[x_2, y_1]\,y_2 + x_1y_1\,[x_2, y_2]. \qquad (10)\end{aligned}$$

Similarly, we get on evaluating in the other way,

$$\begin{aligned}[x_1x_2, y_1y_2] &= [x_1x_2, y_1]\,y_2 + y_1\,[x_1, x_2y_2],\\ &= \{[x_1, y_1]\,x_2 + x_1\,[x_2, y_1]\}\,y_2 + y_1\,\{[x_1, y_2]\,x_2 + x_1\,[x_2, y_2]\},\\ &= [x_1, y_1]\,x_2y_2 + x_1\,[x_2, y_1]\,y_2 + y_1\,[x_1, y_2]\,x_2 + y_1x_1\,[x_2, y_2]. \qquad (11)\end{aligned}$$

Equate the two expressions on the right-hand sides of (10) and (11), and cancel out $x_1\,[x_2, y_1]\,y_2$ and $y_1\,[x_1, y_2]\,x_2$, which occur in both in the same order, then

$$[x_1, y_1]\,y_2x_2 + x_1y_1\,[x_2, y_2] = [x_1, y_1]\,x_2y_2 + y_1x_1\,[x_2, y_2],$$

i.e., on transposing,

$$(x_1y_1 - y_1x_1)\,[x_2, y_2] = [x_1, y_1]\,(x_2y_2 - y_2x_2). \qquad (12)$$

In classical mechanics $xy - yx = 0$, so that each side in (12) vanishes. But this is not the case in quantum mechanics, so that we get

$$\frac{x_1y_1 - y_1x_1}{[x_1, y_1]} = \frac{x_2y_2 - y_2x_2}{[x_2, y_2]} = \text{a constant}. \qquad (13)$$

Evidently, since $\dfrac{x_1y_1 - y_1x_1}{[x_1, y_1]}$ does not depend on x_2, y_2, and $\dfrac{x_2y_2 - y_2x_2}{[x_2, y_2]}$. does not depend on x_1, y_1, the constant value of the ratio cannot depend on x_1, y_1, x_2, y_2. Moreover, it must commute with $x_1y_1 - y_1x_1$ and $x_2y_2 - y_2x_2$, so that it cannot be a matrix but an ordinary number.

By the use of the correspondence principle, Dirac found that for large quantum numbers the value of the constant comes out to be $\frac{ih}{2\pi}$. Generalising this result, Dirac assumed that always

$$xy - yx = \frac{ih}{2\pi}[x, y], \tag{14}$$

i.e., that the difference between the "Heisenberg products" xy and yx of two quantum numbers, is $\frac{ih}{2\pi}$ times their Poisson Bracket expression, where $i = \sqrt{-1}$, and h is Planck's constant. This is Dirac's fundamental new postulate in quantum mechanics. It is based on the correspondence principle, and we see that this principle is woven as an integral part in the new mechanics, and not loosely added to it as in the former theory. The formula (14) is an axiom of the new theory and requires no formal proof, just as Newton's laws of motion require no formal proof in the classical dynamics. All the proof that the formula (14) requires is that it should lead to results in agreement with experiments.

It can be easily verified that the quantum Poisson Brackets defined by (14) satisfy all the conditions (*a*) to (*i*). They are thus uniquely determined.

The equation (14) fills the gap, as it were, left over by the non-commutative multiplication of the quantum variables. It replaces the classical condition that $xy - yx = 0$. Further, it is a natural generalisation of the classical theory, since for the limiting case $h \to 0$, the formula (14) goes over to the classical formula $xy - yx = 0$.

6·5. *The Quantum Conditions and the Equations of Motion.*

In § 2·5 we pointed out that a canonical or contact transformation is one which leaves the form of Hamilton's canonical equations

$$\dot{q}_r = \frac{\partial H}{\partial p_r}, \quad \dot{p}_r = -\frac{\partial H}{\partial q_r}, \quad (r = 1, 2, \cdots, s), \tag{1}$$

unchanged, *i.e.*, if the co-ordinates q_r, p_r are changed canonically to the co-ordinates Q_r, P_r, and H is expressed in Q_r, P_r, then the equations (1) become

$$\dot{Q}_r = \frac{\partial H}{\partial P_r}, \quad \dot{P}_r = -\frac{\partial H}{\partial Q_r}, \quad (r = 1, 2, \cdots, s). \tag{2}$$

We have already mentioned in the last section that the Poisson Brackets of the classical dynamics are invariant under a canonical transformation, so that

$$\begin{aligned}[x, y] &= \sum_{r=1}^{s} \left(\frac{\partial x}{\partial q_r}\frac{\partial y}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial y}{\partial q_r}\right), \\ &= \sum_{r=1}^{s} \left(\frac{\partial x}{\partial Q_r}\frac{\partial y}{\partial P_r} - \frac{\partial x}{\partial P_r}\frac{\partial y}{\partial Q_r}\right).\end{aligned}$$

Consequently, Poisson Brackets are most suited for determining the quantum conditions, and equations of motion in the new quantum mechanics.

From Dirac's equation (14) § 6·4, we have for the generalised co-ordinates q_r, p_r, $(r = 1, 2, \cdots s)$,

$$\left.\begin{aligned} q_m q_n - q_n q_m &= \frac{ih}{2\pi}[q_m, q_n], \\ p_m p_n - p_n p_m &= \frac{ih}{2\pi}[p_m, p_n], \\ q_m p_n - p_n q_m &= \frac{ih}{2\pi}[q_m, p_n], \end{aligned}\right\} \tag{3}$$

for all m and all n. Combining this with the equations (c), (d) and (e) of § 6· 4, we get for all m and all n,

$$\left.\begin{aligned} q_m q_n - q_n q_m &= 0 \\ p_m p_n - p_n p_m &= 0, \\ q_m p_n - p_n q_m &= 0 \text{ if } m \neq n \\ &= \frac{ih}{2\pi} \text{ if } m = n. \end{aligned}\right\} \tag{4}$$

These are the new quantum conditions, which were deduced at first by Heisenberg and Born and others, from the old quantum conditions of Bohr's theory, *viz.*,

$$\oint p_r dq_r = n_r h, \qquad (r = 1, 2, \cdots, s),$$

by a use of the correspondence principle. They are derived here by a method due to Dirac.

Differentiation of a Matrix w.r. to time.—Consider any dynamical variable x of the classical theory which does not contain the time explicitly, so that it depends only on the co-ordinates q_r, p_r. Its derivative with respect to time is

$$\frac{dx}{dt} = \dot{x} = \sum_{r=1}^{s} \left(\frac{\partial x}{\partial q_r}\dot{q}_r + \frac{\partial x}{\partial p_r}\dot{p}_r\right).$$

Substituting in this the values of $\dot{q}_r$, $\dot{p}_r$ from the equations of motion (1), we get

$$\dot{x} = \sum_{r=1}^{s} \left(\frac{\partial x}{\partial q_r}\frac{\partial \mathrm{H}}{\partial p_r} - \frac{\partial x}{\partial p_r}\frac{\partial \mathrm{H}}{\partial q_r}\right)$$
$$= [x, \mathrm{H}].$$

Thus every derivative with respect to time can be expressed as a Poisson Bracket. This result is taken straight into the quantum mechanics, and we assume in general

$$\dot{x} = [x, \mathrm{H}], \tag{5}$$

where H is the Hamiltonian of the dynamical system. Combined with (14) § 6· 4, this gives

$$\frac{ih}{2\pi}\,\dot{x} = \frac{ih}{2\pi}\,[x, \mathrm{H}] = x\mathrm{H} - \mathrm{H}x. \tag{6}$$

For (6) to have a meaning in quantum mechanics, we should give a definition of the differentiation of a matrix with respect to time. We assume that $\dot{x}$ is a matrix whose general mn element is the time derivative of the mn element of the matrix x. Thus, since

$$x = (x_{mn}\, e^{2\pi i\nu_{mn} t}), \tag{7}$$

the time derivative comes out to be

$$\dot{x} = (2\pi i\nu_{mn}\, x_{mn}\, e^{2\pi i\nu_{mn} t}). \tag{8}$$

If x is constant in time, then since $\nu_{mn} \neq 0$ when $m \neq n$, and $\nu_{nn} = 0$, all the elements of x not in the leading diagonal are zero, and those in the leading diagonal are constants, so that all the elements of the matrix $\dot{x}$ are zero, *i.e.*, by definition, the matrix $\dot{x} = 0$. Then from (6) we get $x\mathrm{H} = \mathrm{H}x$, *i.e.*, x commutes with the Hamiltonian H. In this case, therefore, x must be a diagonal matrix.

Writing q_r, p_r in succession for x in (6), we get

$$\left.\begin{aligned} \frac{ih}{2\pi}\,\dot{q}_r &= q_r\mathrm{H} - \mathrm{H}q_r \\ \frac{ih}{2\pi}\,\dot{p}_r &= p_r\mathrm{H} - \mathrm{H}p_r \end{aligned}\right\} (r = 1, 2, \cdots s). \tag{9}$$

These are the equations of motion in quantum mechanics corresponding to the equations (1) in the classical theory. They could have been derived immediately from (8) and (9) of § 6· 4, if we knew the meaning of $\dot{q}_r$, $\dot{p}_r$.

It is not difficult to show that any dynamical variable x which commutes with p_r and with q_r, commutes also with any function f of p_r, q_r. In fact, since $p_r x = x p_r$,

$$0 = p_r x - x p_r = [p_r, x] = \sum_{k=1}^{s} \left(\frac{\partial p_r}{\partial q_k} \frac{\partial x}{\partial p_k} - \frac{\partial p_r}{\partial p_k} \frac{\partial x}{\partial q_k} \right)$$
$$= \frac{\partial x}{\partial q_r} \quad (r = 1, 2, \cdots s).$$

Similarly

$$0 = q_r x - x q_r = [q_r, x] = \sum_{k=1}^{s} \left(\frac{\partial q_r}{\partial q_k} \frac{\partial x}{\partial p_k} - \frac{\partial q_r}{\partial p_k} \frac{\partial x}{\partial q_k} \right)$$
$$= \frac{\partial x}{\partial p_r} \quad (r = 1, 2, \cdots, s).$$

Therefore

$$[f(q_r, p_r), x] = \sum_{k=1}^{s} \left(\frac{\partial f}{\partial q_k} \frac{\partial x}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial x}{\partial q_k} \right) = 0;$$

so that

$$fx - xf = 0, \tag{10}$$

which proves our theorem.

Further, if in (5) we replace x by t, we get:

$$\frac{dt}{dt} = 1 = [t, \mathrm{H}].$$

Comparing this with (5) § 6· 4, we find that the time t and the energy H are a pair of canonical variables.

6· 6. *The Equation of Energy and Bohr's Frequency Relation.*

We know that the Hamiltonian H in classical mechanics denotes the total energy, *i.e.*, the sum of the kinetic energy and the potential energy, and that for dynamical systems under conservative forces this total energy is constant in time. In the last section we have shown that any quantum dynamical variable x which is constant in time must satisfy the equation $\dot{x} = 0$, that is, x must be a diagonal matrix $(x_{mn}\, \delta_{mn})$.

In order to prove the equation of energy, *i.e.*, that H is constant in time, we must show that

$$\dot{\mathrm{H}} = 0, \text{ i.e., } \mathrm{H} = (\mathrm{H}_{mn}\, \delta_{mn}), \tag{1}$$

or that H is a diagonal matrix. This follows immediately, if we substitute H for x in (5) or (6) § 6· 5.

We have seen in § 6· 2 that each of the stationary states of the dynamical system is characterised by a term value of the set

$$T_1, T_2, \cdots, T_n, \cdots; \tag{2}$$

and the Ritz Combination principle says that the frequency of a spectral line due to the transition m to n is given by the relation (3) § 6· 2:

$$\nu_{mn} = T_m - T_n. \tag{3}$$

Now consider the diagonal matrix T given by:

$$T = \begin{pmatrix} T_1 & 0 & 0 & \cdots & \cdots \\ 0 & T_2 & 0 & \cdots & \cdots \\ 0 & 0 & T_3 & 0 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} = (T_{mn}\,\delta_{mn}), \tag{4}$$

where $T_{nn} = T_n$, and $T_{mn} = 0$ when $m \neq n$.

Since both H and T are diagonal matrices, $H - hT$ is also a diagonal matrix, and consequently commutes with every other matrix. Thus, for any dynamical variable $x = (x_{mn}\, e^{2\pi i\nu_{mn}t})$, we have,

$$x\,(H - hT) - (H - hT)\,x = 0, \tag{5}$$

or taking the mn element:

$$\sum_k \{x_{mk}\, e^{2\pi i\nu_{mk}t}\,(H_{kn} - hT_{kn}) - (H_{mk} - hT_{mk})\, x_{kn}\, e^{2\pi i\nu_{kn}t}\} = 0,$$

or

$$x_{mn}\, e^{2\pi i\nu_{mn}t}\,(H_n - hT_n) - (H_m - hT_m)\, x_{mn}\, e^{2\pi i\nu_{mn}t} = 0.$$

Dividing out by $x_{mn}\, e^{2\pi i\nu_{mn}t}$, we find

$$(H_n - hT_n) - (H_m - hT_m) = 0,$$

or, on account of (3):

$$H_m - H_n = h\,(T_m - T_n) = h\nu_{mn}.$$

Thus we get finally:

$$\nu_{mn} = \frac{H_m - H_n}{h}, \tag{6}$$

which is Bohr's frequency relation.

In the last two sections we have considered dynamical systems having s degrees of freedom. The term-values are therefore characterised, not by one index n, but by s indices n_1, n_2, $\cdots$, n_s, thus

$$T_{n_1, n_2, \cdots, n_s}.$$

Each transition is therefore from a stationary state $(m_1, m_2, \cdots, m_s)$ to another state $(n_1, n_2, \cdots, n_s)$. The dynamical variable is represented really by a $2s$-dimensional matrix:

$$x = (x_{n_1, m_2, \cdots, m_s ; n_1, n_2, \cdots, n_s} \, e^{2\pi i \nu_{m_1, m_2, \cdots, m_s ; n_1, n_2, \cdots, n} t}).$$

We have taken a single letter m to represent the whole complex $(m_1, m_2, \cdots, m_s)$. It can easily be seen that $2s$-dimensional matrices have properties analogous to those of the two dimensional matrices.

Moreover, it should be remarked that the quantum mechanical matrices are of an infinite order. In fact, from the quantum condition $qp - pq = \frac{ih}{2\pi}\,\delta$, we get for the nn element:

$$\sum_k (q_{nk}p_{kn} - p_{nk}q_{kn}) = \frac{ih}{2\pi}.$$

Therefore summing for all n, we have

$$\sum_n \sum_k (q_{nk}p_{kn} - p_{nk}q_{kn}) = \sum_n \frac{ih}{2\pi} = n\,\frac{ih}{2\pi}.$$

If n were finite, the left-hand side would be zero, because then the terms would cancel out in pairs. The right-hand side shows that it can never be zero, and therefore n must be infinite.

6·7. *The Harmonic Oscillator.*

As an illustration of the application of quantum mechanics, we shall consider the simple problem of the harmonic oscillator in one dimension. This was actually the problem solved by Heisenberg in his first paper in which he outlined the new mechanics. The details of the theory were supplied by Born and Jordan in September 1925.

To be able to compare the method and results of the new mechanics with those of Bohr's theory, we give first the treatment of the problem according to the older theory.

Suppose that a particle P of mass μ is elastically bound to a centre O, so that the force on P is always directed towards O, and is proportional to the distance q between OP. Then P will oscillate linearly about O and is called a "harmonic oscillator". Such oscillators were first introduced by Planck in his theory of the black body radiation, when he assumed that the black body consisted of harmonic oscillators. If P is an electrified particle, then due to its accelerated motion, it will emit radiation of frequency ν_0.

Let the velocity of the particle be $\dot{q}$, and the momentum $p = \mu\dot{q}$. The kinetic and potential energies are

$$T = \tfrac{1}{2}\mu\dot{q}^2 = \frac{1}{2\mu}p^2, \quad V = \tfrac{1}{2}kq^2, \tag{1}$$

where k is a constant.

The Hamiltonian H is therefore

$$H = T + V = \frac{1}{2\mu}p^2 + \tfrac{1}{2}kq^2. \tag{2}$$

Hamilton's equations of motion are

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{\mu}; \quad \dot{p} = -\frac{\partial H}{\partial q} = -kq. \tag{3}$$

Therefore, differentiating the first equation, we get

$$\ddot{q} = \frac{\dot{p}}{\mu} = \frac{-kq}{\mu}.$$

Now, it is known from classical electrodynamics that

$$\frac{k}{\mu} = (2\pi\nu_0)^2, \tag{4}$$

so that the equation of motion becomes

$$\ddot{q} + (2\pi\nu_0)^2 q = 0. \tag{5}$$

In Bohr's theory, a solution of this equation is

$$q = a\cos(2\pi\,\nu_0 t + \alpha), \tag{6}$$

where a is the amplitude and α the phase. Therefore

$$p = \mu\dot{q} = -\mu 2\pi\,\nu_0\,a\sin(2\pi\nu t + \alpha), \tag{7}$$

so that

$$\begin{aligned} H &= \frac{1}{2\mu}p^2 + \tfrac{1}{2}kq^2 = \tfrac{1}{2}\mu\{\dot{q}^2 + (2\pi\nu_0)^2 q^2\}, \\ &= \tfrac{1}{2}\mu\{4\pi^2\nu_0^2a^2\sin^2(2\pi\nu_0 t + \alpha) + 4\pi^2\nu_0^2a^2\cos^2(2\pi\nu_0 t + \alpha)\}, \\ &= 2\mu\pi^2\nu_0^2a^2. \end{aligned} \tag{8}$$

We shall now find the action variable J.

$$\begin{aligned} J &= \oint p\,dq = \int_0^{\frac{1}{\nu_0}} p\dot{q}\,dt, \\ &= \int_0^{\frac{1}{\nu_0}} \mu 2\pi\nu_0 a\sin(2\pi\nu_0 t + \alpha).\ 2\pi\nu_0 a\sin(2\pi\nu_0 t + \alpha)\,dt, \end{aligned}$$

$$= \mu 4\pi^2\nu_0^2 a^2 \int_0^{\frac{1}{\nu_0}} \sin^2 (2\pi\nu_0 t + \alpha)\, dt,$$

$$= 4\pi^2\mu\nu_0^2 a^2 \cdot \frac{1}{2\pi\nu_0} \cdot \int_0^{2\pi} \sin^2\theta\, d\theta,$$

$$= 4\pi^2\mu\nu_0^2 a^2 \cdot \frac{1}{2\pi\nu_0} \cdot 4 \cdot \int_0^{\frac{\pi}{2}} \sin^2\theta\, d\theta,$$

so that finally

$$\mathrm{J} = 4\pi^2\mu\nu_0^2 a^2 \cdot \frac{1}{2\pi\nu_0} \cdot 4 \cdot \frac{\pi}{4},$$

$$= 2\pi^2\mu\, a^2\nu_0. \tag{9}$$

Comparing (8) and (9) we have

$$\mathrm{H} = \mathrm{J}\nu_0. \tag{10}$$

Now we introduce Bohr's quantum condition

$$\mathrm{J} = nh, \ (n \text{ integer}) \tag{11}$$

so that equation (10) becomes

$$\mathrm{H} = n\,(h\nu_0). \tag{12}$$

Equation (12) shows that the energy of the oscillator is an integral multiple of $h\nu_0$. This was Planck's original postulate introduced in 1900. Later, it was found experimentally that the energy must be a half integral multiple.

We can find the value of the amplitude a from (9) and (11),

$$a^2 = \frac{\mathrm{J}}{2\pi^2\mu\nu_0} = \frac{nh}{2\pi^2\mu\nu_0}, \tag{13}$$

and the displacement q is given by

$$q = \sqrt{\frac{nh}{2\pi^2\mu\nu_0}} \cos (2\pi\nu_0 t + \alpha). \tag{14}$$

The phase α is indeterminable.

Equation (12) is not true. Experiments show that we must have $\mathrm{H} = (n + \frac{1}{2})\, h\nu_0$ instead.

In the new mechanics of Heisenberg, the equations (1) to (5) are valid, but q is a matrix

$$q = (q_{lk}) = (a_{lk}\, e^{2\pi i\nu_{lk}t}), \tag{15}$$

and similarly p is a matrix:

$$p = (p_{lk}) = (b_{lk}\, e^{2\pi i\nu_{lk}t})\,;$$

so that equations (6) to (14) have no meaning in this mechanics.

The equation of motion (5) is true for each one of the elements lk of the matrix (15), therefore, we get

$$(\nu_0^2 - \nu_{lk}^2)\, q_{lk} = 0. \tag{16}$$

From this we see that

$$\text{either } q_{lk} = 0 \text{ or } \nu_{lk} = \pm\, \nu_0. \tag{17}$$

This means that all q_{lk} vanish, except those elements for which $\nu_{lk} = +\,\nu_0$ or $-\,\nu_0$.

The order of numbering the elements lk is at our disposal. We distinguish between the two possibilities $\nu_{lk} = +\,\nu_0$ or $-\,\nu_0$, by supposing that

$$\left.\begin{aligned} \nu_{lk} &= +\,\nu_0 \text{ corresponds to a transition } l \rightarrow l-1, \\ &\quad i.e., \text{ when } k = l-1\,; \\ \nu_{lk} &= -\,\nu_0 \text{ corresponds to a transition } l \rightarrow l+1, \\ &\quad i.e., \text{ when } k = l+1. \end{aligned}\right\} \tag{18}$$

For the elements q_{lk} we get therefore from (17) and (18),

$$\left.\begin{aligned} q_{lk} &= 0 \text{ when } k \neq l \mp 1, \\ q_{lk} &\neq 0 \text{ when } k = l \mp 1. \end{aligned}\right\} \tag{19}$$

Thus the matrix x has the following form:

$$q \equiv \begin{pmatrix} 0 & q_{12} & 0 & 0 & 0 & 0 & 0 \\ q_{21} & 0 & q_{23} & 0 & 0 & 0 & 0 \\ 0 & q_{32} & 0 & q_{34} & 0 & 0 & 0 \\ 0 & 0 & q_{43} & 0 & q_{45} & 0 & 0 \\ 0 & 0 & 0 & q_{54} & 0 & q_{56} & 0 \end{pmatrix}. \tag{20}$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

This is all the information the equation of motion (5) can supply; it cannot give us the value of the non-vanishing elements in (20). For this we must use the quantum condition

$$qp - pq = \frac{ih}{2\pi}\,\delta \text{ or } pq - qp = \frac{h}{2\pi i}\,\delta. \tag{21}$$

On account of $p = \mu\dot{q}$, we have

$$p_{lk} = \mu .\, 2\pi i \nu_{lk} q_{lk}, \tag{22}$$

so that taking the element ll of (21),

$$(pq - qp)_{ll} = 2\pi\mu i \sum_m (\nu_{lm}\, q_{lm}\, q_{ml} - q_{lm}\, \nu_{ml}\, q_{ml}) = \frac{h}{2\pi i}\,;$$

or since $\nu_{lk} = -\,\nu_{lk}$, the last equation becomes

$$\sum_m \nu_{lm}\, q_{lm}\, q_{ml} = -\,\frac{h}{8\pi^2\mu}. \tag{23}$$

On account of the equations (18) and (19), the sum reduces only to two terms, *viz.*, those for which $m = l - 1$, and $m = l + 1$,

$$\nu_{l,l-1}\, q_{l,l-1}\, q_{l-1,l} + \nu_{l,l+1}\, q_{l,l+1}\, q_{l+1,l} = \frac{h}{8\pi^2\mu}.$$

But on account of (18), $\nu_{l,l-1} = +\nu_0$, $\nu_{l,l+1} = -\nu_0$, therefore,

$$q_{l,l-1}\, q_{l-1,l} - q_{l,l+1}\, q_{l+1,l} = -\frac{h}{8\pi^2\mu\nu_0}.$$

But from (7) § 6· 2, we have

$$q_{l-1,l} = \overset{*}{q}_{l,l-1},\quad q_{l,l+1} = \overset{*}{q}_{l+1,l},$$

so that

$$q_{l,l-1}\, q_{l-1,l} = |q_{l,l-1}|^2,\quad q_{l+1,l}\, q_{l,l+1} = |q_{l+1,l}|^2.$$

The last equation therefore reduces to

$$|q_{l+1,l}|^2 - |q_{l,l-1}|^2 = \frac{h}{8\pi^2\mu\nu_0}. \tag{24}$$

We conclude from (24) that the q^2 form an arithmetic series which is unlimited to the right, but must break up on the left, since it can contain only positive terms. At what indices we break up the series is still arbitrary, since we have as yet determined only the relative values of l, k, but not their absolute values. We stipulate now that $q_{1,0}$ shall be the last non-vanishing term of the series, and that $q_{0,-1}$ and all the previous terms shall vanish. Then from (24) we get for $l = 0, 1, 2, \cdots, n, \cdots,$

$$|q_{1,0}|^2 = \frac{h}{8\pi^2\mu\nu_0},\quad |q_{21}|^2 = \frac{2h}{8\pi^2\mu\nu_0},\ \cdots,$$

$$|q_{n,n-1}|^2 = \frac{nh}{8\pi^2\mu\nu_0}; \tag{25}$$

so that for the elements themselves we get

$$\left.\begin{aligned} q_{n,n-1} &= \sqrt{\frac{nh}{8\pi^2\mu\nu_0}}\, e^{2\pi i(\nu_0 t + \beta_{n,n-1})}, \\ q_{n,n+1} &= \sqrt{\frac{(n+1)h}{8\pi^2\mu\nu_0}}\, e^{-2\pi i(\nu_0 t + \beta_{n,n+1})}, \end{aligned}\right\} \tag{26}$$

because on account of (18), $\nu_{n,n-1} = +\nu_0$ and $\nu_{n,n+1} = -\nu_0$.

The quantum mechanics furnishes us therefore with the selection rules (18) and (19) for possible transitions, and also the amplitudes (25) and (26) which give the probabilities of the corresponding transitions, *i.e.*, the intensities of the spectral lines.

We go over now to the calculation of the energy values associated with the stationary states. For this we take the Hamiltonian H given by (2) and (4):

$$\mathrm{H} = \frac{1}{2\mu} p^2 + \tfrac{1}{2} k q^2 = \tfrac{1}{2}\mu \{\dot{q}^2 + (2\pi\nu_0)^2 q^2\}. \tag{27}$$

But $\dot{q}$ is the matrix $(2\pi i \nu_{lk}\, q_{lk})$, so that taking the diagonal element H_{nn}, we get from (27):

$$\begin{aligned} \mathrm{H}_{nn} &= \tfrac{1}{2}\mu \{\sum_m \dot{q}_{nm}\dot{q}_{mn} + (2\mu\nu_0)^2 \sum_m q_{nm}q_{mn}\}, \\ &= \tfrac{1}{2}\mu \{\sum_m (2\pi i)^2 \nu_{nm}q_{nm}\nu_{mn}q_{mn} + (2\pi\nu_0)^2 \sum_m q_{nm}q_{mn}\}. \end{aligned}$$

But $\nu_{mn} = -\nu_{nm}$, and $q_{nm}q_{mn} = |\,q_{nm}\,|^2$, therefore

$$\mathrm{H}_{nn} = 2\pi^2\mu \sum_m \{(\nu_{nm}{}^2 + \nu_0{}^2)\,|\,q_{nm}\,|^2\}. \tag{28}$$

On account of (19), the sum (28) reduces to two terms only, *viz.*, those for which $m = n - 1$ and $m = n + 1$, so that

$$\begin{aligned} \mathrm{H}_{nn} &= 2\pi^2\mu \{(\nu_{n,n-1}{}^2 + \nu_0{}^2)|\,q_{n,n-1}\,|^2 + (\nu_{n,n+1}{}^2 + \nu_0{}^2)\,|\,q_{n,n+1}\,|^2\}, \\ &= 2\pi^2\mu \{([\,+\nu_0]^2 + \nu_0{}^2)\,|\,q_{n,n-1}\,|^2 + ([\,-\nu_0{}^2] + \nu_0{}^2)|\,q_{n,n+1}\,|^2\}, \\ &= 4\pi^2\mu\nu_0{}^2 \{|\,q_{n,n-1}\,|^2 + |\,q_{n,n+1}\,|^2\}. \end{aligned}$$

Substituting in this the values of $|\,q_{n,n-1}\,|^2$ and $|\,q_{n,n+1}\,|^2$ from (25),

$$\begin{aligned} \mathrm{H}_{nn} &= 4\pi^2\mu\nu_0{}^2 \left\{\frac{nh}{8\pi^2\mu\nu_0} + \frac{(n+1)h}{8\,\pi^2\mu\nu_0}\right\}, \\ &= \frac{h\nu_0}{2}\{n + n + 1\}, \\ &= (n + \tfrac{1}{2})\,h\nu_0. \end{aligned} \tag{29}$$

Again, from (27), we get for the general element H_{nm} $(m \neq n)$,

$$\begin{aligned} \mathrm{H}_{nm} &= \tfrac{1}{2}\mu \{\sum_l \dot{q}_{nl}\dot{q}_{lm} + (2\pi\nu_0)^2 \sum_l q_{nl}q_{lm}\}, \\ &= \tfrac{1}{2}\mu \{(2\pi i)^2 \sum_l \nu_{nl}q_{nl}\nu_{lm}q_{lm} + (2\pi\nu_0)^2 \sum_l q_{nl}q_{lm}\}, \\ &= 2\pi^2\mu \sum_l \{(\nu_0{}^2 - \nu_{nl}\nu_{ml})\,q_{nl}q_{lm}\}, \\ &= 2\pi^2\mu \{(\nu_0{}^2 - \nu_{n,n-1}\nu_{n-1,m})\,q_{n,n-1}q_{n-1,m} \\ &\qquad + (\nu_0{}^2 - \nu_{n,n+1}\nu_{n+1,m})\,q_{n,n+1}q_{n+1,m}\}, \end{aligned} \tag{30}$$

where, on account of (19) m can be only $n - 2$ or $n + 2$. If m is neither of these, H_{nm} would vanish, so we have

$$\mathrm{H}_{nm} = 0 \text{ except when } m = n,\ n - 2 \text{ or } n + 2. \tag{31}$$

Now let $m = n - 2$, then from (30)

$$H_{n,n-2} = 2\pi^2\mu \, \{\nu_0{}^2 - \nu_{n,n-1}\nu_{n-1,n-2}) \, q_{n,n-1}q_{n-1,n-2}\}.$$

But, on account of (18) $\nu_{n,n-1} = \nu_{n-1,\,n-2} = \nu_0$, so that

$$H_{n,n-2} = 2\pi^2\mu \, \{(\nu_0{}^2 - \nu_0{}^2) \, q_{n,n-1}q_{n-1,n-2}\} = 0. \qquad (32)$$

If $m = n + 2$, then from (30)

$$H_{n,n+2} = 2\pi^2\mu \, \{\nu_0{}^2 - \nu_{n,n+1}\nu_{n+1,n+2}) \, q_{n,n+1}q_{n+1,n+2}\}.$$

But, on account of (18), $\nu_{n,n+1} = \nu_{n+1,n+2} = -\nu_0$, so that

$$H_{n,n+2} = 2\pi^2\mu \, \{(\nu_0{}^2 - \nu_0{}^2) \, q_{n,n+1}q_{n+1,n+2}\} = 0. \qquad (33)$$

Combining (31), (32) and (33) we get

$$H_{nm} = 0 \; (m \neq n). \qquad (34)$$

From (29) and (34) we find, as was to be expected, that H is a diagonal matrix, *i.e.*, H is constant in time, verifying the law of conservation of energy.

The diagonal elements $H_{nn} = H_n$ give therefore the energy values associated with the stationary states, so that:

$$H_n = (n + \tfrac{1}{2}) \, h\nu_0, \quad (n = 0, 1, 2, \cdots). \qquad (35)$$

This formula was first derived by Heisenberg, who showed that the energy values of the harmonic oscillators are half integral multiples of $h\nu_0$, as required by experiment. Bohr's theory had led to the wrong result (12). This was the first brilliant confirmation of Heisenberg's mechanics.

Moreover, from the kinetic theory, we know that the internal energy (the heat energy) of a body is due to the oscillations of the molecules in the body about their equilibrium positions. Experiments with low temperatures had led the German physicist W. Nernst to the conclusion that even at the absolute zero of temperature the body has still a finite energy, different from zero. This hypothesis of Nernst, put forward in 1913, is now the universally accepted third law of Thermodynamics. The absolute zero of temperature is that stationary state for which the quantum number $n = 0$. Now Bohr's theory cannot account for Nernst's law because it gives zero value for the lowest energy of the atom. But Heisenberg's theory gives the value $\frac{1}{2}\, h\nu_0$ for the lowest energy ($n = 0$) of the atom, which accounts for the third law of thermodynamics. In fact, the value $\frac{1}{2}\, h\nu_0$ was being freely used by Planck, Nernst and others in statistical theory, long before the advent of Heisenberg's mechanics.

This example serves to show the coherence and power of Heisenberg's method, which is based on the calculus of matrices. It supplies us by a uniform method the selection rules, the intensities and the energy values for various problems. The problems of the hydrogen spectrum, the Stark and Zeeman-effects, the anomalous Zeeman-effects, etc., were solved successfully on Heisenberg's theory by various writers. They led to the same results in cases where Bohr's theory was known to give correct results. But in cases where Bohr's theory gave wrong results or failed completely, Heisenberg's theory gave results in agreement with experiments.

However, even in this, the simplest, example of the harmonic oscillator, we see that, though the method is straightforward and elementary in principle, the actual calculations are long and tedious. The situation is much worse in other problems mentioned above. As a matter of fact, quantum mechanics would not have made such rapid progress, had it not received impetus from another source.

A few months after the publication of Heisenberg's paper, in January 1926 to be exact, E. Schrödinger put forward his wave-mechanics, based on the ideas of A. Einstein and L. de Broglie about the nature of radiation and matter. We shall give an account of wave-mechanics in the following chapters.

6· 8 (1). *The Angular Momentum.*

Suppose a moving particle has co-ordinates x, y, z, and linear momenta p_x, p_y, p_z. Then its angular momentum vector $\overrightarrow{m}$, with components m_x, m_y, m_z, is defined, as in the classical theory, by the relations :

$$m_x = yp_z - zp_y, \quad m_y = zp_x - xp_z, \quad m_z = xp_y - yp_x. \qquad (1)$$

The angular momentum thus defined satisfies certain identities which can be proved by using the quantum conditions :

$$[x, y] = [y, z] = [z, x] = 0,$$

$$[x, p_x] = [y, p_y] = [z, p_z] = 1,$$

$$[x, p_y] = [x, p_z] = y, p_x] = [y, p_z] = [z, p_x] = [z, p_y] = 0. \quad (2)$$

First of all we have the identity

$$m_x x + m_y y + m_z z = 0. \tag{3}$$

Then we find,

$$\begin{aligned}[m_x, y] &= [yp_z = zp_y, y] = [yp_z, y] - [zp_y, y] \\ &= [y, y]\, p_z + y\, [p_z, y] - [z, y]\, p_y - z\, [p_y, y] = +z.\end{aligned}$$

Thus we have the following identities:

$$\left.\begin{array}{lll} [m_x, y] = 0, & [m_x, y] = z, & [m_x, z] = -y\,; \\ [m_y, x] = -z, & [m_y, y] = 0, & [m_y, z] = x\,; \\ [m_z, x] = y, & [m_z, y] = -x, & [m_z, z] = 0. \end{array}\right\} \tag{4}$$

Similarly we have

$$\begin{aligned}[m_y, p_y] &= [yp_z - zp_y, p_y] = [yp_z, p_y] - [zp_y, p_y] \\ &= [y, p_y]\, p_z + y\, [p_z, p_y] - [z, p_y]\, p_y - z\, [p_y, p_y] = p_z.\end{aligned}$$

Thus

$$\left.\begin{array}{lll} [m_x, p_x] = 0, & [m_x, p_y] = p_z, & [m_x, p_z] = -p_y\,; \\ [m_y, p_x] = -p_z, & [m_y, p_y] = 0. & [m_y, p_z] = p_x\,; \\ [m_z, p_x] = p_y, & [m_z, p_y] = -p_x, & [m_z, p_z] = 0. \end{array}\right\} \tag{5}$$

Further we have

$$\begin{aligned}[m_x, m_y] &= [m_x, zp_x - xp_z] = [m_x, zp_x] - [m_x, xp_z] \\ &= [m_x, z]\, p_x + z\, [m_x, p_x] - [m_x, x]\, p_z - x\, [m_x, p_z] \\ &= -yp_x + xp_y = m_z.\end{aligned}$$

Thus

$$[m_x, m_y] = m_z,\ [m_y, m_z] = m_x,\ [m_z, m_x] = m_y. \tag{6}$$

From these identities, we find that

$$\begin{aligned}[m_x, x^2 + y^2 + z^2] &= x\, [m_x, x] + [m_x, x]\, x + y\, [m_x, y] \\ &\qquad + [m_x, y]\, y + z\, [m_x, z] + [m_x, z]\, z = 0.\end{aligned}$$

Thus

$$[m_x, x^2 + y^2 + z^2] = [m_y, x^2 + y^2 + z^2] = [m_z, x^2 + y^2 + z^2] = 0. \tag{7}$$

Again

$$\begin{aligned}[m_x, p_x^2 + p_y^2 + p_z^2] &= p_x\, [m_x, p_x] + [m_x, p_x]\, p_x + p_y\, [m_x, p_y] \\ &\quad + [m_x, p_y]\, p_y + p_z\, [m_x, p_z] + [m_x, p_z]\, p_z = 0,\end{aligned}$$

so that we have

$$\begin{aligned}[m_x, p_x^2 + p_y^2 + p_z^2] &= [m_y, p_x^2 + p_y^2 + p_z^2] \\ &= [m_z, p_x^2 + p_y^2 + p_z^2] = 0.\end{aligned} \tag{8}$$

The identities (7) and (8) show that each of m_x, m_y, m_z commutes with $x^2 + y^2 + z^2$, and also with $p_x^2 + p_y^2 + p_z^2$.

Therefore each of them would commute with any function of $(x^2 + y^2 + z^2)$ and $(p_x^2 + p_y^2 + p_z^2)$, as proved in the theorem at the end of § 6·5.

Moreover,

$$\begin{aligned}[m_x, m_x^2 + m_y^2 + m_z^2] &= [m_x, m_y^2 + m_z^2] \\ &= m_y\,[m_x, m_y] + [m_x, m_y]\,m_y, \\ &\quad + m_z\,[m_x, m_z] + [m_x, m_z]\,m_z \\ &= 0.\end{aligned}$$

Thus, since $m^2 = m_x^2 + m_y^2 + m_z^2$, we get

$$[m_x, m^2] = [m_y, m^2] = [m_z, m^2] = 0, \tag{9}$$

showing that each component of $\vec{m}$ commutes with m^2.

If there are several particles, with angular momenta $\vec{m}_1$, $\vec{m}_2, \cdots, \vec{m}_n$, each of them will satisfy the identities (1) to (9). The total angular momentum vector $\vec{\mathrm{M}}$ is defined as

$$\vec{\mathrm{M}} = \sum_{r=1}^{n} \vec{m}_r, \tag{10}$$

with the components

$$\mathrm{M}_x = \sum_r m_{rx},\ \mathrm{M}_y = \sum_r m_{ry},\ \mathrm{M}_z = \sum_r m_{rz}.$$

Then, apart from the quantum conditions of the type (2), we have further conditions for $r \neq s$, $(r, s = 1, 2, \cdots, n)$:

$$[x_r, x_s] = 0, \text{ etc.}, [p_{xr}, p_{xs}] = 0, \text{ etc.}, [p_{xr}, q_{xs}] = 0, \text{ etc.} \tag{11}$$

It is evident therefore that any function of the co-ordinates and momenta of one particle commutes with any function of those of another particle.

Now, we have for any particle k,

$$\begin{aligned}[\mathrm{M}_x, y_k] &= [\textstyle\sum_r m_{rx}, y_k] = \sum_r [m_{rx}, y_k] \\ &= [m_{kx}, y_k], \text{ since } [m_{rx}, y_k] = 0 \text{ if } r \neq k, \\ &= z_{kx}, \text{ on account of (4).}\end{aligned}$$

Thus we get for $k = 1, 2, \cdots, n$:

$$\left.\begin{aligned}[\mathrm{M}_x, x_k] &= 0, & [\mathrm{M}_x, y_k] &= z_k, & [\mathrm{M}_x, z_k] &= -y_k; \\ [\mathrm{M}_y, x_k] &= -z_k, & [\mathrm{M}_y, y_k] &= 0, & [\mathrm{M}_y, z_k] &= x_k; \\ [\mathrm{M}_z, x_k] &= y_k, & [\mathrm{M}_z, y_k] &= -x_k, & [\mathrm{M}_z, z_k] &= 0.\end{aligned}\right\} \tag{12}$$

Similarly, corresponding to the relation (5), we obtain for each $k = 1, 2, \cdots, n$:

$$\left.\begin{aligned}[\mathrm{M}_x, p_{kx}] &= 0, & [\mathrm{M}_x, p_{ky}] &= p_{kz}, & [\mathrm{M}_x, p_{kz}] &= -p_{ky}; \\ [\mathrm{M}_y, p_{kx}] &= -p_{kz}, & [\mathrm{M}_y, p_{ky}] &= 0, & [\mathrm{M}_y, p_{kz}] &= p_{kx}; \\ [\mathrm{M}_z, p_{kx}] &= p_{ky}, & [\mathrm{M}_z, p_{ky}] &= -p_{kx}, & [\mathrm{M}_z, p_{kz}] &= 0,\end{aligned}\right\} \tag{13}$$

Further, we find

$$\begin{aligned}[\mathrm{M}_x, \mathrm{M}_y] &= [\sum_r m_{rx}, \sum_s m_{sy}] = \sum_r \sum_s [m_{rx}, m_{sy}] \\ &= \sum_r [m_{rx}, m_{ry}], \text{ since } [m_{rx}, m_{sy}] = 0 \text{ if } s \neq r, \\ &= \sum_r m_{rz} = \mathrm{M}_z, \text{ on account of (6).}\end{aligned}$$

Thus, corresponding to the relation (6) for a single particle we obtain for the angular momentum of a system :

$$[\mathrm{M}_x, \mathrm{M}_y] = \mathrm{M}_z, [\mathrm{M}_y, \mathrm{M}_z] = \mathrm{M}_x, [\mathrm{M}_z, \mathrm{M}_x] = \mathrm{M}_y. \tag{14}$$

From (14) we get by a reasoning similar to (9) if we write $\mathrm{M}^2 = \mathrm{M}_x^2 + \mathrm{M}_y^2 + \mathrm{M}_z^2$,

$$[\mathrm{M}_x, \mathrm{M}^2] = [\mathrm{M}_y, \mathrm{M}^2] = [\mathrm{M}_z, \mathrm{M}^2] = 0. \tag{15}$$

We see that the angular momentum identities hold not only for a single particle, but also for a system of particles.

6· 8 (2). *Selection and Polarisation Rules.*

Suppose that a non-degenerate atomic system is acted upon by forces which have a symmetry about an axis which we shall take to be the z-axis. Then the forces will have no moment round the z-axis, so that from the principle of angular momentum we shall get

$$\frac{d\mathrm{M}_z}{dt} = 0,$$

showing that M_z is a diagonal matrix.

Then, from the relation $[\mathrm{M}_z, z] = 0$, we obtain

$$M_z z - z\mathrm{M}_z = 0,$$

so that

$$\sum_j \{\mathrm{M}_z(kj)\, z(jl) - z(kj)\, \mathrm{M}_z(jl)\} = 0,$$

or, since M_z is a diagonal matrix,

$$\mathrm{M}_z(kk)\, z(kl) - z(kl)\, \mathrm{M}_z(ll) = 0,$$

i.e.,

$$z(kl)\{\mathrm{M}_z(kk) - \mathrm{M}_z(ll)\} = 0. \tag{2}$$

Similarly, from the other two relations $[\mathrm{M}_z, x] = y$, $[\mathrm{M}_z, y] = -x$, or $\mathrm{M}_z x - x\mathrm{M}_z = \frac{ih}{2\pi} y$ and $\mathrm{M}_z y - y\mathrm{M}_z = -\frac{ih}{2\pi} x$, we get

$$\left.\begin{aligned} x(kl)\{\mathrm{M}_z(kk) - \mathrm{M}_z(ll)\} &= \frac{ih}{2\pi} y(kl), \\ y(kl)\{\mathrm{M}_z(kk) - \mathrm{M}_z(ll)\} &= -\frac{ih}{2\pi} x(kl). \end{aligned}\right\} \tag{3}$$

From (3) it appears that, if in a transition $k \rightarrow l$, $M_z\ (kk)$ does not change, then $x\ (kl) = 0$ and $y\ (kl) = 0$, so that the light is polarised parallel to the z-axis.

But if $M_z\ (kk)$ does change, *i.e.*, if $M\ (kk) \neq M_z\ (ll)$, then from (2) we see that $z\ (kl) = 0$.

Now suppose that M_z is increased by an amount Δ in a transition $k \rightarrow l$, *i.e.*, $\Delta = M_z\ (ll) - M_z\ (kk)$, then from (3),

$$x\,(kl)\,\Delta = -\frac{ih}{2\pi}\,y\,(k, l),\ y\,(kl)\,\Delta = \frac{ih}{2\pi}\,x\,(kl).$$

$$\left.\begin{aligned} \{x\,(kl) + iy\,(kl)\}\left(\Delta + \frac{h}{2\pi}\right) &= 0,\\ \{x\,(kl) - iy\,(kl)\}\left(\Delta - \frac{h}{2\pi}\right) &= 0. \end{aligned}\right\} \tag{4}$$

Thus we must either have

$$\left.\begin{aligned} \Delta - \frac{h}{2\pi} &= 0 \text{ and } x\,(kl) + iy\,(kl) = 0,\\ \text{or}\qquad \Delta + \frac{h}{2\pi} &= 0 \text{ and } x\,(kl) - iy\,(kl) = 0. \end{aligned}\right\} \tag{5}$$

This shows that if $M_z\ (kk)$ changes, it can only do so by an amount $\pm \frac{h}{2\pi}$; in each case the light is circularly polarised in the plane xy.

We have thus proved that in general $M_z\ (kk)$ can only change by an amount 0 or $\pm \frac{h}{2\pi}$, so that $M_z\ (kk)$ must be of the form

$$M_z\,(kk) = \frac{h}{2\pi}\ (j + \text{constant}), \tag{6}$$

where j is a positive or negative integer or zero. We see further that j can only change by 0 or $\pm$ 1, so that

$$\Delta j = 0,\ +1,\ -1. \tag{7}$$

$\Delta j = 0$ corresponds to polarisation parallel to the field, and $\Delta j = \pm 1$ to polarisation at right angles to the field.

6·9 (1). *Further Development of Matrix Algebra.*

Adjoint Matrix.—From a matrix a we form another matrix $a^\dagger$ by interchanging the columns and rows, and replacing the elements by their complex conjugates:

$$a = \begin{pmatrix} a_{11} & a_{12} & a_{13}, & \cdots \\ a_{21} & a_{22} & a_{23}, & \cdots \\ a_{31} & a_{32} & a_{33}, & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},\quad a^\dagger = \begin{pmatrix} a_{11}^* & a_{21}^* & a_{31}^* & \cdots \\ a_{12}^* & a_{22}^* & a_{32}^* & \cdots \\ a_{13}^* & a_{23}^* & a_{33}^* & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

i.e.,

$$a^\dagger_{mn} = a^*_{nm}, \quad (m, n = 1, 2, \cdots). \tag{1}$$

The matrix $a^\dagger$ is said to be "adjoint" to the matrix a.

From this definition we immediately deduce that

$$(a + b)^\dagger = a^\dagger + b^\dagger, \tag{2}$$

$$(ab)^\dagger = b^\dagger a^\dagger. \tag{3}$$

The relation (2) can be easily established. To prove (3), we remark that for all m, n, we have

$$(ab)^\dagger_{mn} = (ab)^*_{nm} = (\sum_k a_{nk}\, b_{km})^* = \sum_k a^*_{nk}\, b^*_{km},$$

$$(b^\dagger\, a^\dagger)_{mn} = \sum_k b^\dagger_{mk}\, a^\dagger_{kn} = \sum_k b^*_{km}\, a^*_{nk} = (ab)^\dagger_{mn}.$$

Hermitian Matrix.—As remarked in § 6· 3, if an adjoint matrix $a^\dagger$ is equal to the original matrix a, *i.e.*, if

$$a^\dagger = a, \text{ or } a_{mn}{}^\dagger = a_{mn} = a^*_{nm}, \tag{4}$$

then the matrix a is said to hermitian, or real or self-adjoint. We know that the quantum mechanical matrices must be hermitian, if they are to represent real quantities. The sum of two hermitian matrices is itself hermitian, because from $a^\dagger = a$, and $b^\dagger = b$, we get $(a + b)^\dagger = a^\dagger + b^\dagger = a + b$. But the product of two hermitian matrices is not in general hermitian, because $(ab)^\dagger = b^\dagger a^\dagger = ba$ is not in general equal to ab. Thus the product of two hermitian matrices a and b is only then hermitian when the two matrices are commutative.

Every matrix a can be expressed in the form

$$a = a_1 + ia_2, \tag{5}$$

where a_1 and a_2 are hermitian matrices. For this it is only necessary to put

$$a_1 = \tfrac{1}{2}\,(a + a^\dagger), \quad a_2 = \frac{1}{2i}\,(a - a^\dagger). \tag{6}$$

It is evident that the expression (5) is unique.

THEOREM 1. If a is a hermitian matrix, and S is any arbitrary matrix, then

$$b = \mathrm{S}^\dagger\, a\mathrm{S} \tag{7}$$

is also hermitian.

For we have, since $(\mathrm{S}^\dagger)^\dagger = \mathrm{S}$,

$$b^\dagger = (\mathrm{S}^\dagger a\mathrm{S})^\dagger = \{\mathrm{S}^\dagger\,(a\mathrm{S})\}^\dagger = (a\mathrm{S})^\dagger\,(\mathrm{S}^\dagger)^\dagger$$

$$= \mathrm{S}^\dagger\, a^\dagger\mathrm{S} = \mathrm{S}^\dagger a\mathrm{S} = b. \tag{8}$$

Unitary Matrix.—A matrix which is identical with the inverse of its adjoint matrix is said to be "unitary". Thus if u is a unitary matrix, then

$$u = (u^\dagger)^{-1}. \tag{9}$$

From (9), we deduce on taking inverse of both sides, and remarking that $(a^{-1})^{-1} = a$,

$$u^{-1} = u^\dagger. \tag{10}$$

Thus a unitary matrix can also be defined as that matrix whose inverse is equal to the adjoint matrix. From (10) we obtain further

$$u\, u^\dagger = u^\dagger u = \delta, \tag{11}$$

where δ is the unit matrix.

If we write $v = u^{-1}$, then we find

$$v^\dagger = (u^{-1})^\dagger = (u^\dagger)^\dagger = u,$$

so that

$$(v^\dagger)^{-1} = u^{-1} = v, \tag{12}$$

showing that $v = u^{-1}$ is also a unitary matrix.

THEOREM 2. The product of two unitary matrices is itself unitary.

For, if u_1, u_2 are two unitary matrices, then

$$(u_1 u_2)^\dagger = u_2^\dagger\, u_1^\dagger = u_2^{-1}\, u_1^{-1} = (u_1 u_2)^{-1}. \tag{13}$$

THEOREM 3. If a is an arbitrary matrix, and u a unitary matrix, and if we define

$$b = u^{-1}\, a u, \tag{14}$$

then

$$b^\dagger = u^{-1} a^\dagger u. \tag{15}$$

For, from (14) we find

$$b^\dagger = (u^{-1} a u)^\dagger = u^\dagger\, a^\dagger\, (u^{-1})^\dagger = u^{-1}\, a^\dagger u.$$

In particular, if a is hermitian, we see that $b^\dagger = b$, that is to say, that b is also hermitian. Thus we get the following theorem.

THEOREM 4. If a is a hermitian, and u a unitary, matrix, then $b = u^{-1}\, au$ is also a hermitian matrix.

THEOREM 5. If a and b are two hermitian matrices, and S a matrix such that

$$b = S^{-1} a S, \text{ or } Sb = aS, \tag{16}$$

then the matrices

$$c = SS^\dagger a, \text{ and } d = S^\dagger S b, \tag{17}$$

are also hermitian.

From (16) we have, since $a^\dagger = a$,

$$b^\dagger = (S^{-1}aS)^\dagger = S^\dagger a\,(S^{-1})^\dagger = S^\dagger\,a\,(S^\dagger)^{-1}, \tag{18}$$

where we have used $(S^{-1})^\dagger = (S^\dagger)^{-1}$, *i.e.*, that for any matrix the inverse of its adjoint is equal to the adjoint of its inverse.

But $b^\dagger = b$, therefore from (16) and (18) we get

$$S^\dagger a\,(S^\dagger)^{-1} = S^{-1}a\,S,$$

or

$$SS^\dagger\,a\,(S^\dagger)^{-1} = SS^{-1}\,aS = aS,$$

or

$$SS^\dagger\,a\,(S^\dagger)^{-1}\,S^\dagger = a\,SS^\dagger,$$

or

$$SS^\dagger a = a\,SS^\dagger = (SS^\dagger a)^\dagger, \tag{19}$$

showing that c is a hermitian matrix.

From (19) we get,

$$S^{-1}SS^\dagger a = S^{-1}\,aSS^\dagger,$$

or

$$S^\dagger a = bS^\dagger. \tag{20}$$

Then

$$S^\dagger aS = bS^\dagger S.$$

But we have from (16) that $aS = Sb$, therefore,

$$S^\dagger Sb = bS^\dagger S = (S^\dagger Sb)^\dagger, \tag{21}$$

which proves that d is also a hermitian matrix.

6·9 (2). *Matrix Analysis.*

Matrix-functions.—From the matrices $x_1, x_2, \cdots, x_s$, which can be considered as variables, we can build up polynomials by a finite number of multiplications and additions. Thus

$$P\,(x_1,\cdots,x_s) = \sum_r \sum_{n_1,\cdots,n_r} c^{(r)}_{n_1\cdots n_r}\,x_{n_1}\,x_{n_2}\cdots x_{n_r}, \tag{1}$$

where the c's are numerical coefficients, and $n_1, n_2, \cdots, n_r$ is any arrangement of the numbers $1, 2, \cdots, s$ with repetitions, and r denotes the degree of the term.

We define the "adjoint polynomial" by

$$P^\dagger\,(x_1,\cdots,x_s) = \sum_r \sum_{n_1,\cdots,n_r} c^{(r)*}_{n_1\cdots n_r}\,x_{n_r}\cdots x_{n_2}\,x_{n_1}. \tag{2}$$

From (2) we find easily:

$$\begin{aligned} P^\dagger\,(x_1{}^\dagger,\cdots,x_s{}^\dagger) &= \sum_r \sum_{n_1,\cdots,n_r} c^{(r)*}_{n_1\cdots n_r}\,x_{n_r}{}^\dagger\cdots x_{n_2}{}^\dagger\,x_{n_1}{}^\dagger \\ &= \{P\,(x_1,\cdots,x_s)\}^\dagger. \end{aligned} \tag{3}$$

Thus we see that the adjoint polynomial (2) is characterised by the fact that it is the adjoint matrix of the original polynomial $P(x_1, \cdots, x_s)$, when in $P^\dagger(x_1, \cdots, x_s)$ the matrix-variables x are replaced by $x^\dagger$.

The polynomial $P(x_1, \cdots, x_s)$ is called "self-adjoint," when

$$P^\dagger(x_1, \cdots, x_s) = P(x_1, \cdots, x_s), \tag{4}$$

holds identically in $x_1, \cdots, x_s$.

If the variables x are hermitian, *i.e.*, $x^\dagger = x$, then we get, on account of (3) and (4),

$$\begin{aligned}\{P(x_1, \cdots, x_s)\}^\dagger &= P^\dagger(x_1{}^\dagger, \cdots, x_s{}^\dagger) = P^\dagger(x_1, \cdots, x_s) \\ &= P(x_1, \cdots, x_s),\end{aligned} \tag{5}$$

showing that a polynomial in hermitian variables is itself hermitian.

Starting from polynomials, we can proceed in the limit to infinite power series, provided they converge, and thus define general matrix functions :

$$f(x_1, \cdots, x_s).$$

THEOREM 1. If x_1 and x_2 are two matrices, and if

$$y_1 = S^{-1} x_1 S, \qquad y_2 = S^{-1} x_2 S, \tag{6}$$

where S is any arbitrary matrix, then

$$\left.\begin{aligned} y_1 + y_2 &= S^{-1}(x_1 + x_2) S, \\ y_1 y_2 &= S^{-1}(x_1 x_2) S. \end{aligned}\right\} \tag{7}$$

For, we have

$$y_1 + y_2 = S^{-1} x_1 S + S^{-1} x_2 S = S^{-1}(x_1 + x_2) S,$$

since the distributive law holds for matrices. Similarly

$$\begin{aligned} y_1 y_2 &= (S^{-1} x_1 S)(S^{-1} x_2 S) = S^{-1} x_1 (S S^{-1}) x_2 S \\ &= S^{-1} x_1 \delta x_2 S = S^{-1} x_1 x_2 S, \end{aligned}$$

since the associative law also holds for the matrices.

For convenience, we have taken only two matrices x_1 and x_2 in Theorem 1. Obviously, the theorem holds for any finite number of matrices, and for any polynomials. Proceeding to the limit, we conclude that the theorem would hold for any power series, or for any function which can be expressed as a power series. Thus we have the following general theorem.

THEOREM 2. If in a matrix-function $f(x_1, x_2, \cdots, x_s)$, the arguments $x_1, \cdots, x_s$ are replaced by new matrix-variables $y_1, y_2, \cdots, y_s$ given by

$$y_1 = S^{-1} x_1 S, \; y_2 = S^{-1} x_2 S, \; \cdots, \; y_s = S^{-1} x_s S, \tag{8}$$

where S is any arbitrary matrix, then the function $f(x_1, \cdots, x_s)$ is transformed into the function

$$f(y_1, \cdots y_s) = S^{-1} f(x_1, \cdots, x_s) S. \tag{9}$$

Now let the arguments $x_1, \cdots, x_s$ be commutative, *i.e.*, let

$$x_k x_l = x_l x_k, \; (k, l = 1, 2, \cdots, s),$$

and let the matrix-function $f(x_1, \cdots, x_s)$ be defined by the power series,

$$f(x_1, \cdots, x_s) = \sum_{n_1, \cdots, n_s} c_{n_1 \cdots n_s} x_1^{n_1} \cdots x_2^{n_2} \cdots x_s^{n_s}. \tag{10}$$

If the coefficients c are real, then from (2) we get

$$\begin{aligned} f^\dagger(x_1, \cdots, x_s) &= \sum_{n_1, \cdots, n_s} c_{n_1 \cdots n_s} x_s^{n_s} \cdots x_2^{n_2} x_1^{n_1} \\ &= \sum_{n_1, \cdots, n_s} c_{n_1 \cdots n_s} x_1^{n_1} x_2^{n_2} \cdots x_s^{n_s} \\ &= f(x_1, \cdots, x_s), \end{aligned} \tag{11}$$

showing that f is self-adjoint.

In particular, we see that every function of a single variable with real coefficients, *viz.*,

$$f(x) = \sum_n c_n x^n, \tag{12}$$

is self-adjoint. For example, the exponential function

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{13}$$

is self-adjoint.

From (13) we deduce immediately that e^0 is the unit matrix δ.

Moreover, if x_1 and x_2 are commutative matrices, then from (13) we find, as in the ordinary analysis,

$$e^{x_1 + x_2} = e^{x_1} e^{x_2}. \tag{14}$$

Therefore, since $-x$ always commutes with x, we get

$$e^x e^{-x} = e^{x-x} = e^0 = \delta,$$

showing that e^{-x} is the inverse of e^x:

$$(e^x)^{-1} = e^{-x}. \tag{15}$$

Further, if a is a hermitian matrix, then on account of (5) we conclude that e^a is also hermitian.

On the other hand, e^{ia} will be a unitary matrix, for

$$(e^{ia})^{\dagger} = e^{-ia\dagger} = e^{-ia}, \quad \text{(since } a^{\dagger} = a\text{)},$$
$$= (e^{ia})^{-1}, \text{ on account of (15).} \tag{16}$$

It can be shown that, conversely, each unitary matrix u can be expressed in the form e^{ia} with the help of a hermitian matrix a.

Matrix-differentiation.—If $f(x, y, z, \cdots)$ is a matrix-function of the independent matrix-variables $x, y, z, \cdots$, then we define the partial derivative of f with respect to x by the relation

$$\frac{\partial f}{\partial x} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \{f(x + \epsilon\delta, y, z, \cdots) - f(x, y, z,)\}, \tag{17}$$

where ϵ is an ordinary number.

From this definition we find immediately:

$$\frac{dx}{dx} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \{x + \epsilon\delta - x\} = \delta;$$
$$\frac{dx^2}{dx} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \{(x + \epsilon\delta)^2 - x^2\},$$

so that

$$\left(\frac{dx^2}{dx}\right)_{mn} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \sum_k \{(x_{mk} + \epsilon\delta_{mk})(x_{kn} + \epsilon\delta_{kn}) - x_{mk}x_{kn}\}$$
$$= \sum_k (x_{mk}\delta_{kn} + \delta_{mk}x_{kn}) = 2x_{mn}.$$

Thus, we find

$$\frac{dx^2}{dx} = 2x. \tag{18}$$

We find in general the rules for differentiating the sum and the product of two matrix-functions $f(x, y, z, \cdots)$ and $g(x, y, z, \cdots)$:

$$\frac{\partial}{\partial x}(f + g) = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}, \tag{19}$$

and

$$\frac{\partial}{\partial x}(fg) = \frac{\partial f}{\partial x} g + f \frac{\partial g}{\partial x}. \tag{20}$$

The rule (19) is obvious from definition. To prove (20) we see that

$$f(x + \epsilon\delta, y, z, \cdots)\, g(x + \epsilon\delta, y, z, \cdots)$$
$$- f(x, y, \cdots)\, g(x, y, z, \cdots)$$
$$= f(x + \epsilon\delta, y, z, \cdots) \{g(x + \epsilon\delta, y, z, \cdots) - g(x, y, z, \cdots)\}$$
$$+ \{f(x + \epsilon\delta, y, z, \cdots) - f(x, y, z, \cdots)\}\, g(x, y, z, \cdots).$$

Dividing both sides of this equation by ϵ and proceeding to the limit $\epsilon \longrightarrow 0$, we obtain (20).

The rules (19) and (20) can be extended to any finite number of functions. Thus we get

$$\frac{dx^n}{dx} = nx^{n-1}, \tag{21}$$

and

$$\frac{de^x}{dx} = e^x. \tag{22}$$

6· 9 (3). *Transformation of a Matrix to a Diagonal Matrix. Determination of Eigenvalues.*

We shall first illustrate the method by a particular example, in which we take the matrix to be of rank 2 :

$$a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{1}$$

Let S be the transformation matrix, and d the diagonal matrix

$$d = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}, \tag{2}$$

to which a is transformed by means of the transformations

$$S^{-1} a S = d, \textit{ i.e.}, a S = S d. \tag{3}$$

Taking the m, n component of both sides, we obtain

$$\sum_{k=1}^{2} a_{mk} S_{kn} = \sum_{k=1}^{2} S_{mk} d_{kn} = S_{mn} d_n,$$

or

$$a_{m1} S_{1n} + a_{m2} S_{2n} = S_{mn} d_n, \quad (m, n = 1, 2). \tag{4}$$

We wish to find d_1 first, so that we set $n = 1$ in (4). Then for $m = 1, 2$ we get the two equations

$$a_{11} S_{11} + a_{12} S_{12} = S_{11} d_1,$$
$$a_{11} S_{11} + a_{22} S_{21} = S_{21} d_1,$$

from which eliminating S_{11} and S_{21}, we obtain

$$\begin{vmatrix} a_{11} - d_1 & a_{12} \\ a_{21} & a_{22} - d_1 \end{vmatrix} = 0. \tag{5}$$

In order to find d_2, we put $n = 2$ in (4) ; then for $m = 1, 2$, we get the two equations

$$a_{11} S_{12} + a_{12} S_{22} = S_{12} d_2,$$
$$a_{21} S_{12} + a_{22} S_{22} = S_{22} d_2.$$

Eliminating S_{12} and S_{22} from these two equations, we obtain

$$\begin{vmatrix} a_{11} - d_2 & a_{12} \\ a_{21} & a_{22} - d_2 \end{vmatrix} = 0. \tag{6}$$

From (5) and (6) we see that d_1 and d_2 are the two roots of the equation in λ:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0. \tag{7}$$

This is called a secular equation, and as proved in algebra, it has always real and distinct roots, provided that a is a hermitian matrix. Thus to find the diagonal matrix d, and consequently the eigenvalues of a, we have only to solve the equation (7).

Generally, let a be a hermitian matrix of rank r:

$$a = (a_{mn}), \ (m, n = 1, 2, \cdots, r), \tag{8}$$

and let it be desired to transform it to a diagonal matrix d:

$$d = (d_m \, \delta_{mn}), \tag{9}$$

by means of the transformation

$$S^{-1} a S = d, \ i.e., \ a S = S d. \tag{10}$$

Taking the mn component of both sides, we get

$$\sum_{k=1}^{r} a_{mk} S_{kn} = \sum_{k=1}^{r} S_{mk} d_{kn} = S_{mn} d_n. \tag{11}$$

Suppose we wish to determine any particular d_n, say d_1. Then writing $n = 1$ in (11) and putting $m = 1, 2, \cdots, r$ successively, we obtain:

$$\begin{aligned} &(a_{11} - d_1) S_{11} + a_{12} S_{21} + a_{13} S_{31} + \cdots + a_{1r} S_{r1} = 0, \\ &a_{21} S_{11} + (a_{22} - d_1) S_{21} + a_{23} S_{31} + \cdots + a_{2r} S_{r1} = 0, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &a_{r1} S_{11} + a_{r2} S_{21} + \cdots \qquad + (a_{rr} - d_1) S_{r1} = 0. \end{aligned} \tag{12}$$

Eliminating $S_{11}, S_{21}, \cdots, S_{r1}$ from these equations, we find

$$\begin{vmatrix} a_{11} - d_1 & a_{12} & a_{13} \cdots a_{1r} \\ a_{21} & a_{22} - d_1 & a_{23} \cdots a_{2r} \\ \cdot \ \cdot \ \cdot & \cdot \ \cdot \ \cdot & \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \\ a_{r1} & a_{r2} & a_{r3} \cdots a_{rr} - d_1 \end{vmatrix} = 0. \tag{13}$$

If we had wished to determine any other d_n we would have obtained the same equation (13), except that d_1 would have been replaced by d_2, d_3, $\cdots$, d_r respectively. We conclude therefore that the eigenvalues d_1, $d_2, \cdots$, d_r of the matrix a, are the roots of the equation

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} & \cdots & a_{rr} - \lambda \end{vmatrix} = 0. \qquad (14)$$

These roots are real, since the matrix a is assumed to be hermitian.

It will be noticed that this procedure of determining the eigenvalues of a matrix, is the same as that of reducing a homogeneous quadratic expression to its normal form, or that of reducing the equation of a second degree surface to the standard form with reference to the principal axes.

It has been proved by Hellinger and Tœplitz that this procedure for determining the eigenvalues can be extended to infinite matrices, with the difference, however, that in the latter case a continuous "spectrum" (range of eigenvalues) is also obtained along with the discrete set.

6· 9 (4). *Canonical Transformations.*

Let q_r, p_r $(r = 1, 2, \cdots, n)$, be the co-ordinates and momenta of any dynamical system, satisfying the quantum conditions

$$q_r q_s - q_s q_r = 0,\ p_r p_s - p_s p_r = 0,$$
$$q_r p_s - p_s q_r = \frac{ih}{2\pi}\,\delta. \qquad (1)$$

We transform to new variables Q_r, P_r by means of the transformations

$$Q_r = S^{-1} q_r S,\ P_r = S^{-1} p_x\, S\ (r = 1, 2, \cdots, n), \qquad (2)$$

where S is any matrix. If $f\,(q_r, p_r)$ is any function of the original variables q_r, p_r, and $f\,(Q_r, P_r)$ the same function of the new variables Q_r, P_r, then we know from Theorem 2 of § 6· 9 (2), that

$$f\,(Q_r, P_r) = S^{-1} f\,(q_r, p_r)\, S. \qquad (3)$$

On account of (1), we find therefore,

$$Q_r Q_s - Q_s Q_r = S^{-1}\, O\, S = 0,$$
$$P_r P_s - P_s P_r = S^{-1}\, O\, S = 0,$$
$$Q_r P_s - P_s Q_r = S^{-1} \frac{ih}{2\pi}\, \delta S = \frac{ih}{2\pi}\, S^{-1}\, S = \frac{ih}{2\pi}\, \delta. \qquad (4)$$

The transformations (2) therefore preserve the form of the quantum conditions, and are therefore canonical.

If H (q_r, p_r) is the Hamiltonian of the dynamical system, then it is transformed to

$$H(Q_r, P_r)) = S^{-1} H(q_r, p_r) S. \tag{5}$$

If we wish to find the energy levels of the dynamical system, we shall have to calculate the eigenvalues of the Hamiltonian H (q_r, p_r). As we have shown in the last section, the discrete set of eigenvalues of H is found by reducing the Hamiltonian H to a diagonal matrix.

In actual dynamical systems, the variables q_r, p_r are hermitian. The new variables Q_r, P_r, given by the transformations (2) must also be hermitian. But in Theorem 1 of § 6· 9 (1), we proved that if a is any hermitian matrix, and S any arbitrary matrix, then $S^\dagger a S$ is also hermitian. Thus, if we wish to make a canonical transformation which should preserve the hermitian character as well, we must take a transformation matrix S such that it should satisfy the two conditions:

$$\left.\begin{aligned} Q &= S^{-1} q S, \\ Q &= S^\dagger q S, \end{aligned}\right\} \tag{6}$$

i.e., we must take S such that

$$S^\dagger = S^{-1}. \tag{7}$$

This means that S must be a unitary matrix. Thus we find that all transformation matrices in quantum mechanics are unitary matrices.

CHAPTER VII.

THE WAVE THEORY OF MATTER.

7·1. *De Broglie's Conception of Matter-Waves.*

IN the last chapter we pointed out that from the year 1923 onwards it was becoming increasingly apparent that the quantum theory of the atom as given by Bohr was not completely true. In fact, several workers were looking at that time for a revision of ideas which would lead to a better theory. Heisenberg's 'Matrix Theory,' published in July 1925, was the first alternative which was more or less complete, and which was able to give quantitative results in the simple instances of the harmonic and anharmonic oscillators, and the hydrogen atom.

Actually, however, another point of view which broke away, not only from Bohr's theory, but from every other theory of matter conceived in the whole history of Science, was proposed already in January 1924 by the French physicist Louis de Broglie. But de Broglie's ideas were not developed far enough by him to lead to any quantitative results, and not much notice was taken of them until Schrödinger revived them in 1926.

De Broglie was led to his theory by a consideration of the quantum theory of light proposed by Einstein, and applied successfully to explain the phenomena of the photo-electric effect, the black-body radiation and the Compton effect. He argued that as it has been found necessary to introduce a corpuscular conception along with the wave-conception in the theory of light, the opposite step should also be taken and the wave-concept introduced along with the corpuscular concept in the theory of matter. To the dualism in the conception of radiation would then correspond an analogous dualism in the conception of matter. De Broglie's fundamental contribution is the idea of waves which are associated with the moving particles of matter.

He assumed that "each individual particle was connected with a system of 'phase waves' in such a way that the path of this particle coincided with a ray of the corresponding wave-system." The waves, as it were, carry along the particles and hence each particle must possess its own train of waves.

Thus if the energy of the particle is ϵ, and its mass m, then the particle has, associated with itself, a system of phase waves whose frequency ν is given by Einstein's equation:

$$\epsilon = mc^2 = h\nu, \tag{1}$$

where h is Planck's constant.

7 · 2. *De Broglie's Theory.*

Generalising Einstein's ideas about the corpuscular nature of radiation, we assume that there is a complete correspondence between the corpuscular and the wave theories even in the case of material particles. The only difference in this case is that for material particles the rest mass m_0 is not zero.

Then corresponding to the energy ϵ (or, what is the same thing, the mass m) and the momentum $p = mv$ of the particle in the corpuscular conception, we shall have the frequency ν and the wave-length λ in the wave-conception, and the relation between these two sets of quantities will be that given by (17) and (18) of § 4 · 2:

$$p = mv = hk = \frac{h}{\lambda}, \tag{1}$$

and

$$\epsilon = mc^2 = h\nu, \tag{2}$$

where k is the " wave-number " defined by

$$k = \frac{1}{\lambda}, \tag{3}$$

and

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{4}$$

Now, we have

$$p^2 = m^2v^2 = m_0{}^2 \frac{v^2}{1 - \frac{v^2}{c^2}} = m_0c^2 \cdot \frac{\frac{v^2}{c^2}}{1 - \frac{v^2}{c^2}};$$

and

$$\frac{\epsilon^2}{c^2} = m^2c^2 = m_0{}^2c^2 \cdot \frac{1}{1 - \frac{v^2}{c^2}},$$

so that

$$\frac{\epsilon^2}{c^2} - p^2 = \frac{m_0c^2}{1 - \frac{v^2}{c^2}}\left(1 - \frac{v^2}{c^2}\right),$$

or

$$\frac{\epsilon^2}{c^2} - p^2 = m_0^2c^2. \tag{5}$$

Substituting $p = hk$ and $\epsilon = h\nu$ from (1) and (2) in (5):

$$\frac{h^2\nu^2}{c^2} - h^2k^2 = m_0^2c^2,$$

or

$$m_0^2 = \frac{h^2}{c^2}\left(\frac{\nu^2}{c^2} - k^2\right). \tag{6}$$

Equation (6) gives us the value of the rest mass m_0 when ν and k are known.

For a photon we know that the velocity of propagation is c, so that

$$c = \nu\lambda = \frac{\nu}{k},$$

and therefore on account of (6) $m_0 = 0$, showing that the rest mass of a photon is zero, as required.

But, for a material particle $m_0 \neq 0$, and the velocity of propagation of the phase waves is u given by the relation

$$u = \nu\lambda = \frac{\nu}{k}. \tag{7}$$

We can determine u from (6) in terms of m_0 and λ. From (6) we get

$$\frac{\nu^2}{c^2} - k^2 = \frac{m_0^2c^2}{h^2},$$

or

$$\frac{\nu^2}{k^2c^2} - 1 = \frac{m_0^2c^2}{h^2k^2} = \frac{m_0^2c^2}{h^2}\lambda^2,$$

or

$$\frac{\nu^2}{k^2c^2} = 1 + \frac{m_0^2c^2}{h^2}\lambda^2,$$

so that

$$\frac{\nu^2}{k^2} = c^2\left(1 + \frac{m_0^2c^2}{h^2}\lambda^2\right). \tag{8}$$

The velocity u of the propagation of phase-waves is, therefore,

$$u = \frac{\nu}{k} = c\sqrt{1 + \frac{m_0^2c^2}{h^2}\lambda^2}. \tag{9}$$

For a material particle $m_0 > 0$, and we see from (9) that $u > c$. This is, however, no violation of the theory of relativity, because according to that theory the velocity of no *material* particle can exceed c, or even be equal to c. The velocity u is only the velocity of propagation of the phase-waves associated with the particle, and as given by (9) it is always greater than c.

We can easily find a relation between u and v, the latter being the velocity of the material particle. From (1) and (2) we get

$$\lambda = \frac{h}{p} = \frac{h\sqrt{1 - \frac{v^2}{c^2}}}{m_0 v}, \tag{10}$$

and

$$\nu = \frac{\epsilon}{h} = \frac{mc^2}{h} = \frac{m_0 c^2}{h\sqrt{1 - \frac{v^2}{c^2}}}. \tag{11}$$

Therefore

$$u = \nu\lambda = \frac{m_0 c^2}{h\sqrt{1 - \frac{v^2}{c^2}}} \times \frac{h\sqrt{1 - \frac{v^3}{c^2}}}{m_0 v},$$

or

$$u = \frac{c^2}{v}. \tag{12}$$

The equation (12) also shows that the velocity u of the propagation of phase-waves is much greater than c, because the velocity v of the material particle is much smaller than the velocity c of light.

From (12) we deduce that if $v = 0$, then $u = \infty$; that is, in the frame of reference which is rigidly attached to the electron, and in which the electron is therefore at rest, the phase-waves would travel with infinite velocity. But, since $u = \nu\lambda$, the wave-length in this frame of reference would therefore be infinite, and the wave would consequently have the same phase throughout space.

The wave-function in this frame has therefore the most simple form

$$\psi' = \psi_0' \cos 2\pi \nu_0 t', \tag{13}$$

where ν_0 is the frequency and t' the time in the frame of reference in which the particle is at rest. Now, let the particle be moving with a velocity v in the direction of the x-axis relative to an observer S. Then from the Lorentz transformation we have:

$$t' = \frac{t - \dfrac{xv}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \tag{14}$$

For the observer S the wave-function would therefore be:

$$\psi = \psi_0 \cos 2\pi \frac{\nu_0}{\sqrt{1 - \dfrac{v^2}{c^2}}} \left(t - \frac{x}{\dfrac{c^2}{v}} \right),$$

or

$$\psi = \psi_0 \cos 2\pi\nu \left(t - \frac{x}{u} \right), \tag{15}$$

where

$$\nu = \frac{\nu_0}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \tag{16}$$

Formula (16) gives the law of transformation of frequency for different frames of reference, which is the same as the law of transformation of mass.

The equation (15) shows once again that the wave travels in the x-direction with a velocity $u = \dfrac{c^2}{v}$.

It is often simpler to express the wave-function ψ as an exponential function. For (15) we write therefore

$$\psi = \psi_0 \, e^{-2\pi i\nu \left(t - \frac{x}{u} \right)}.$$

We wish to introduce here the corpuscular quantities E, p, instead of the wave-quantities ν, u. Writing $\nu = \mathrm{E}/h$, (16) becomes:

$$\psi = \psi_0 \, e^{\frac{2\pi i}{h} \left(\frac{\mathrm{E}}{u} x - \mathrm{E}t \right)}. \tag{17}$$

But we have, on account of $\mathrm{E} = mc^2$ and $u = \dfrac{c^2}{v}$,

$$\frac{\mathrm{E}}{u} = \frac{mc^2}{u} = mv = p, \tag{18}$$

so that (17) becomes finally

$$\psi = \psi_0 \, e^{\frac{2\pi i}{h} (px - \mathrm{E}t)}. \tag{19}$$

7 · 3. *Wave-Group and Group Velocity.*

For the sake of convenience we consider a wave composed of a superposition of two components of equal amplitudes but slightly different frequencies and phase-velocities, *i.e.*, a superposition of two waves associated with particles having slightly different energies and momenta. The wave-function is then

$$\psi = \psi_0\, e^{\frac{2\pi i}{h}(px - \mathrm{E}t)} + \psi_0\, e^{\frac{2\pi i}{h}\{(p + \Delta p)x - (\mathrm{E} + \Delta \mathrm{E})t\}}$$

$$= \psi_0 \left\{1 + e^{\frac{2\pi i}{h}(\Delta p.\, x - \Delta \mathrm{E}.\, t)}\right\} e^{\frac{2\pi i}{h}(px - \mathrm{E}t)} \quad . \qquad (1)$$

The resultant wave is therefore the original wave $e^{\frac{2\pi i}{h}(px - \mathrm{E}t)}$, with a variable amplitude

$$\psi_0 \left\{1 + e^{\frac{2\pi i}{h}(\Delta p.\, x - \Delta \mathrm{E}.\, t)}\right\} . \qquad (2)$$

This means that the form of the group continually alters as it advances. The velocity with which any modulation of the group, such as the maximum of amplitude, is propagated is called the " group-velocity ".

From (2) it is evident that initially when $t = 0$, the amplitude has the maximum value $2\,\psi_0$ at the point $x = 0$. Suppose this maximum value of the amplitude occurs again after a time t_1 at the point $x = x_1$, then we must have

$$\Delta p \cdot x_1 - \Delta \mathrm{E} \cdot t_1 = 0,$$

or

$$\frac{x_1}{t_1} = \frac{\Delta \mathrm{E}}{\Delta p}. \qquad (3)$$

The maximum of the amplitude has travelled a distance x_1 in time t_1, so that according to definition the group-velocity w is given by

$$w = \frac{x_1}{t_1} = \frac{\Delta \mathrm{E}}{\Delta p}. \quad . \qquad (4)$$

But $\qquad \Delta \mathrm{E} = \Delta(\tfrac{1}{2}mv^2 + \text{const.}) = mv\,\Delta v,$

and $\qquad \Delta p = \Delta(mv) = m\,\Delta v,$

giving $\qquad \dfrac{\Delta \mathrm{E}}{\Delta p} = \dfrac{mv \cdot \Delta v}{m \cdot \Delta v} = v. \qquad (5)$

From (4) and (5) we conclude that

$$w = v, \tag{6}$$

i.e., the group-velocity w of the phase-waves is equal to the velocity v of the material particle.

It appears therefore that the material particle is surrounded by a group of waves, the group moving with the same velocity v as the particle, while the individual waves move with a velocity $u = \frac{c^2}{v}$.

7 · 4. *The Analogy between Mechanics and Optics.*

De Broglie's ideas described in the last three sections were developed further by E. Schrödinger in January 1926. Schrödinger revived Hamilton's conceptions about the analogy existing between mechanics and optics.

It is easy to see that there is a complete correspondence between the laws of geometrical or ray-optics and those of the classical mechanics governing the motion of a particle in a conservative field of force, inasmuch as Fermat's principle of least time in optics, and Maupertius' principle of least action in mechanics have a common basis.

Fermat's principle in optics states that the actual ray between two fixed points A and B is such that the time which light takes to go from A to B is a minimum. We can express this condition analytically. Let the refractive index at any point (x, y, z) of the medium be $n\ (x, y, z)$, and the velocity at that point be $u\ (x, y, z)$, then from the definition of the refractive index we have :

$$n = \frac{c}{u} \text{ or } u = \frac{c}{n}, \tag{1}$$

where c is the velocity of light in vacuum.

The time that a light ray takes to travel an element ds of its path would be $\frac{ds}{u} = \frac{n}{c}\, ds$. Therefore Fermat's principle states that $\int_A^B \frac{n}{c}\, ds$ is a minimum, or that

$$\delta \int_A^B \frac{n}{c}\, ds = 0. \tag{2}$$

Similarly, let a particle of mass m move in a field of force given by the potential V (x, y, z). As the field is conservative, the total energy $E = T + V$ is conserved, where T is the kinetic energy. Then

$$T = \tfrac{1}{2} mv^2 = E - V. \tag{3}$$

From (21) § 1 · 1, we have Maupertius' principle of least action that

$$\delta \int_{t_1}^{t_2} 2T \, dt = 0,$$

or

$$\delta \int_{t_1}^{t_2} mv^2 \, dt = 0,$$

or

$$\delta \int_{t_1}^{t_2} mv \frac{ds}{dt} \, dt = 0,$$

or

$$\delta \int_{A}^{B} mv \, ds = 0.$$

But from (3) we have

$$m^2v^2 = 2m (E - V),$$

so that the principle of least action becomes

$$\delta \int_{A}^{B} \sqrt{2m} \, \sqrt{E - V} \, ds = 0. \tag{4}$$

By a comparison of (2) and (4) we see that the path of a material particle is the same as the trajectory of a light ray in a medium whose refractive index is proportional to $\sqrt{E - V}$.

Schrödinger pursued this analogy between mechanics and optics to its logical conclusion. He argued that there are gross and fine phenomena both in the domain of optics and in the domain of mechanics. The gross or large-scale phenomena in optics are the phenomena of rectilinear propagation and of reflection and refraction, and these can be dealt with adequately by geometrical or ray optic which is a sort of a corpuscular theory.

Similarly, for the large-scale phenomena of the material bodies, *i.e.*, for phenomena coming under the head of macro-mechanics, the particle mechanics developed by Newton, Lagrange, Hamilton, Jacobi and finally by Einstein, is quite adequate.

But, for the finer phenomena of physical optics, such as interference and diffraction, ray-optics or the corpuscular theory is quite useless, and we must have recourse to a wave-analysis.

Exactly in the same way, for the microphenomena taking place on an atomic scale, such as the motion of individual electrons, atoms or molecules, the particle mechanics is totally inadequate, and we must use a wave theory.

7 · 5. *Schrödinger's Wave Equation.*

Consider the differential equation of the wave optics:

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}=\frac{1}{u^2}\frac{\partial^2\phi}{\partial t^2}, \tag{1}$$

where u is the phase-velocity of light in the medium. Since ϕ is periodic in time, we can write

$$\phi\,(x, y, z, t)=\psi\,(x, y, z)\,e^{2\pi i\nu t}, \tag{2}$$

where ν is the vibration frequency.

Substituting (2) in (1), and dividing out by $e^{2\pi i\nu t}$, we get

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}+4\pi^2\frac{\nu^2}{u^2}\psi=0. \tag{2'}$$

From (7) § 7 · 2, we know that if λ is the wave-length, then $u=\nu\lambda$, *i.e.*, $\frac{\nu}{u}=\frac{1}{\lambda}$, so that on writing this in (2′), we obtain:

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}+\frac{4\pi^2}{\lambda^2}\psi=0. \tag{3}$$

Now let the wave be a de Broglie wave associated with a particle of mass m, moving with a velocity v (x, y, z) in a conservative field of force having a potential V (x, y, z). Then from equation (1) § 7 · 2, we know that the wave-length λ of a de Broglie wave is given by

$$\lambda=\frac{h}{\text{momentum}}=\frac{h}{mv},$$

so that

$$\frac{1}{\lambda^2}=\frac{m^2v^2}{h^2}, \tag{4}$$

showing that the wave-length of a de Broglie wave varies continuously from point to point. But, if T is the kinetic energy, and E the total energy, we have

$$\tfrac{1}{2}mv^2=\mathrm{T}=\mathrm{E}-\mathrm{V},$$

or

$$m^2v^2 = 2m\,(E - V). \tag{5}$$

Combining (4) and (5) we get,

$$\frac{1}{\lambda^2} = \frac{2m\,(E - V)}{h^2}. \tag{6}$$

Substituting (6) into (3) we have

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{8\pi^2 m}{h^2}(E - V)\,\psi = 0. \tag{7}$$

This is the famous wave-equation of Schrödinger, which is derived here by a simple and direct method, being perhaps also the most significant physically. The actual derivation by Schrödinger was along completely different lines. In a later section, we shall deduce de Broglie's theory as well as Schrödinger's wave-equation from Heisenberg's equation (4) § 6 · 5.

The essence of wave mechanics is that the motion of a material particle is governed by the second order wave-equation (7), and not by the first order differential equation of Hamilton and Jacobi, as in classical mechanics.

7 · 6. *Experiments on the Diffraction of Electrons.*

When de Broglie put forward his conception of material waves associated with a moving particle, it was only a hypothesis complementary to Einstein's idea of the quantum of light. In Schrödinger's hands this hypothesis of the electron-waves led to results in perfect agreement with experiments. It was thus demonstrated that the wave-theory of matter was not inconsistent with the experimental facts. Soon afterwards, a direct proof of the wave-like behaviour of electrons was forthcoming.

The first observation of the diffraction effect of the electrons at crystals was made by C. J. Davisson and C. H. Kunsman in 1923, though they had no idea at that time as to the significance of their results, namely that they had to do with the diffraction of electrons. They were making experiments to determine the distribution-in-angle of electrons, scattered by a target of ordinary nickel. In these experiments they found that the intensity of the reflected electrons depended to a large extent on the angle of reflection. It was also observed that the convergence of intensity in certain privileged directions depended strongly on the velocities of the electrons. During a continuation of these investigations, they were quite accidentally made aware of the fact that the structure of the crystals had a profound influence on the results. The experimental tube was broken at a time when the target was at a high temperature. This resulted in the target being heavily oxidized by the inrushing air. The oxide was eventually reduced, and a layer of the target removed by vaporization but only after prolonged heating at various high temperatures in hydrogen and in vacuum.

When the experiments were continued, it was found that the distribution-in-angle of the scattered electrons had been completely changed. This marked alteration in the scattering pattern was traced to a recrystallization of the target that occurred during the prolonged heating. Before the accident the target consisted of many small crystals, but after the accident it was converted into a few large ones.

It seemed probable from this result that the intensity of scattering from a single crystal would exhibit a marked dependence on crystal direction, and the authors set about making experiments for an investigation of this dependence. Even at this stage they had not conceived the idea of electron-diffraction.

In 1925, however, W. Elsasser proposed that according to de Broglie's theory of material waves, the results of Davisson and Kunsman should be interpreted as the diffraction of electrons. He argued that on de Broglie's theory the electron is considered as a group of waves whose wave-length is given by the formula $\lambda = \frac{h}{mv} = \frac{h\sqrt{1 - \frac{v^2}{c^2}}}{m_0 v}$, where m_0 is the rest-mass and v the velocity of the electron. For electrons of 25,000 volts energy, the wave-length λ calculated from the above formula is about $0{\cdot}75 \times 10^{-6}$ cm. This is of the order of that of hard X-rays, and the waves associated with electrons of this energy should behave in many respects like hard X-rays. In fact, on de Broglie's view, there is a close analogy between the two. The quantum effects of X-rays are regarded as due to centres of energy guided by waves, while in the case of electrons, the motion of the electric charge, in which the energy is centred, is regarded as taking place along the rays of the group of phase-waves associated with it. In particular, the electron should show diffraction effects when passed through a crystal identical with those shown by X-rays of the same wave-length.

Inspired by this suggestion, Davisson and Germer tried to confirm it by direct experiments. In April 1927, they showed experimentally that very similar phenomena are observed if the incident electron beam were replaced by a beam of mono-chromatic X-rays of adjustable wave-lengths. They projected a stream of electrons on to a crystal of nickel, and found that electrons were reflected from the surface of the crystal, just as X-rays are reflected in von Laue's experiments. The intensity of the reflected pencil of the electrons was measured as a function of the direction of reflection. The experiment showed that, in general, the velocity of the elcetrons remained unaltered or changed very slightly, and that the electrons were diffusely reflected in all directions. For certain directions of reflections and for certain electronic velocities, the reflected pencils had an exceptional intensity. These special directions of reflection are found experimentally to be the same as those obtained in the diffraction of X-rays, provided only that the wave-length of the X-rays is $\lambda = \frac{h}{mv}$, where m is the mass and v the velocity of the electron.

They concluded that because of these similarities between the scattering of electrons by the crystal and the scattering of X-rays by 3- and 2-dimensional gratings, it is not only possible but most simple and natural to believe that the electrons are diffracted at a crystal exactly as though they behaved like waves of wave-length $\lambda = \frac{h}{mv}$.

A little later, *i.e.*, in November of the same year, G. P. Thomson advanced the matter a stage further, demonstrating by means of other series of experiments that a wave process is associated with moving electrons. In essence, his experiment consisted in sending a fine beam of approximately homogeneous cathode rays through a very thin celluloid film at normal incidence. According to classical physics, the electrons would emerge on the other side of the film as a disordered mob, each individual moving with a different speed in a different direction. But the result of G. P. Thomson's experiment was quite the reverse.

When the electrons were received on a photographic plate, Thomson found that the central spot formed by the undeflected rays is surrounded by rings recalling in appearance the haloes formed by mist round the sun. This pattern is identical with that obtained in a Debye-Scherrer apparatus with X-rays of the same wave-length. This showed that the passage of the electrons through the interstices of the metal is like the passage of waves.

In 1930 and succeeding years, O. Stern and his collaborators discovered that atomic and molecular rays of hydrogen and helium also show diffraction phenomena when they are reflected at the surface of crystals.

The dual nature, *i.e.*, the corpuscular and wave nature, of both matter and radiation is an experimentally established fact, and all future physical theories would have to embody it as one of the most fundamental data of our knowledge of the universe.

CHAPTER VIII.

WAVE MECHANICS.

8·1. *Eigen-values and Eigen-functions.*

CONSIDER the differential equation

$$\frac{d^2y}{dx^2} + \lambda y = 0, \tag{1}$$

and suppose we want to find a regular solution $y\ (x)$, *i.e.*, a solution such that y, $\frac{dy}{dx}, \frac{d^2y}{dx^2}$, are continuous, in the whole domain

$$0 \leqslant x \leqslant \pi. \tag{2}$$

Suppose also that y satisfies the boundary conditions

$$y\ (0) = 0, \tag{3}$$

and

$$y\ (\pi) = 0. \tag{4}$$

λ is a parameter which does not depend on x. The most general solution of (1) is

$$y = a \sin \sqrt{\lambda}\ x + b \cos \sqrt{\lambda}\ x. \tag{5}$$

But the condition (3) requires that $b = 0$. The solution therefore reduces to

$$y = a \sin \sqrt{\lambda}\ x, \tag{6}$$

where a cannot be zero, if the solution is not to become trivial.

The condition (4) now requires that

$$\sin \sqrt{\lambda}\ \pi = 0, \tag{7}$$

and this is the case only when $\sqrt{\lambda}$ is an integer $n \geqslant 1$. Therefore we must have

$$\lambda = n^2\ (n = 1, 2, 3, \cdot \cdot \cdot \infty). \tag{8}$$

We see thus that the boundary problem in general has no solution except when the parameter λ has one of a discrete set of values. These values $1^2, 2^2, 3^2, 4^2, \cdot \cdot \cdot$ of λ, for which alone the boundary problem has a solution, are called the "Eigen-values" (proper-values, characteristic values) of the differential equation, and the corresponding solutions

$$y_n = a_n \sin nx\ (n = 1, 2, 3, \cdot \cdot \cdot), \tag{9}$$

are called the "Eigen-functions" (proper functions, characteristic functions) of the equation.

Every differential equation which contains a parameter, and which satisfies certain boundary conditions possesses a set of eigen-values and eigen-functions.

The wave-equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{8\pi^2 m}{h^2} (\mathrm{E} - \mathrm{V})\, \psi = 0, \tag{10}$$

has also a parameter E, which, being the total energy, does not depend on x, y, z. As ψ is a physical quantity, it cannot be infinite anywhere, and we must therefore determine a solution of the equation (10) which satisfies the condition that it remains finite, continuous and single valued throughout the whole space. As we shall see in concrete examples, this will be the case only when E has a certain discrete set of values, or a certain continuous range of values. The discrete set E_n of the eigen-values constitutes the energy levels of the atomic system. We can thus determine the energy levels quite naturally without finding it necessary to have recourse to any arbitrary assumptions like the quantum conditions. The great merit of Schrödinger's contribution lies in having noticed for the first time that the problem of quantisation can be reduced to a problem of determining the eigen-values of differential equations.

8 · 2 *The Polynomial Method for Determining the Eigen-values.*

Having defined the eigen-values in the last section, we describe here one or two simple methods of determining them. For the sake of convenience as well as definiteness, we illustrate these methods by selecting well-known equations for discussion.

The first method is the so-called " polynomial method," or " the method of terminating series ".

Consider the differential equation

$$(1 - x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + \left(\lambda - \frac{m^2}{1 - x_2}\right) y = 0, \tag{1}$$

where λ is the parameter and m is any positive constant. The conditions to be satisfied by the solution $y(x)$ is that it must be finite, continuous and single-valued for all values of x from $-\infty$ to $+\infty$.

If we divide the equation (1) by $(1 - x^2)$, we get

$$\frac{d^2y}{dx^2} - \frac{2x}{1 - x^2}\frac{dy}{dx} + \left(\frac{\lambda}{1 - x^2} - \frac{m^2}{(1 - x^2)^2}\right)y = 0. \tag{2}$$

From equation (2) we see that for $x = -1$, $x = +1$ or $x = \infty$, the coefficients of $\frac{dy}{dx}$ and of y become infinite. Such values of the independent variable x, for which one of the coefficients becomes infinite, provided that the coefficient of $\frac{d^2y}{dx^2}$ is made unity after division, are called "the singular points" of the linear differential equation. Thus the singular points of (2) are $x = -1$, $x = +1$ and $x = \infty$.

A singular point is called a "pole" when there is possible an expansion in power series in which the terms with negative powers are finite in number. Otherwise the singular point is an "essential singular point". The singular points of (2), *viz.*, -1, $+1$ and $+\infty$ are all poles.

A value of x for which none of the coefficients becomes infinite, is called an "ordinary point" of the differential equation.

For an ordinary point or a pole, $x = x_0$, the solution which is valid in its neighbourhood can be determined as a power series in $z = x - x_0$:

$$y = z^{\alpha}\,(a_0 + a_1 z = a_2 z^2 + \cdots \infty), \tag{3}$$

where α is a real number to be determined and the a's depend on the parameter λ.

For the equation (2), we consider first the pole $x_0 = 1$, and set therefore $z = x - 1$ or $x = z + 1$. Then

$$\frac{dy}{dx} = \frac{dy}{dz}\cdot\frac{dz}{dx} = \frac{dy}{dz}.$$

Substituting $x = z + 1$ and $\frac{dy}{dx} = \frac{dy}{dz}$ in (2), we get

$$\frac{d^2y}{dz^2} + \frac{2(z + 1)}{z(z + 2)}\frac{dy}{dz} - \left(\frac{\lambda}{z\,(z + 2)} + \frac{m^2}{z^2\,(z + 2)^2}\right)y = 0. \tag{4}$$

From the series (3), we find:

$$\frac{d^2y}{dz^2} = a_0\,\alpha\,(\alpha - 1)\,z^{\alpha-2} + a_1\,(\alpha + 1)\alpha\,z^{\alpha-1} + \cdots$$
$$+ a_n\,(\alpha + n)\,(\alpha + n - 1)\,z^{\alpha+n-2} + \cdots,$$

$$\frac{dy}{dz} = a_0\alpha\ z^{\alpha-1} + a_1(\alpha+1)\,z^{\alpha} + \cdots + a_n(\alpha+n)\,z^{\alpha+n-1} + \cdots$$

$$y = a_0\,z^{\alpha} + a_1\,z^{\alpha+1} + \cdots + a_n\,z^{\alpha+n} + \cdots$$

Moreover, we have from the binomial theorem :

$$\frac{2(z+1)}{z(z+2)} = \frac{2(z+1)}{z}\,\frac{1}{2\left(1+\frac{z}{2}\right)} = \frac{1}{z}(z+1)\left(1+\frac{z}{2}\right)^{-1}$$

$$= \frac{1}{z}(1+z)\left(1-\frac{z}{2}+\frac{z^2}{4}-\cdots\right)$$

$$= \frac{1}{z} + \tfrac{1}{2} + \text{positive powers of } z.$$

$$-\frac{\lambda}{z(z+2)} = -\frac{\lambda}{z}\cdot\frac{1}{2\left(1+\frac{z}{2}\right)} = -\frac{\lambda}{2z}\left(1+\frac{z}{2}\right)^{-1}$$

$$= -\frac{\lambda}{2z}\left(1-\frac{z}{2}+\frac{z^2}{4}-\cdots\right)$$

$$= -\frac{\lambda}{2z} + \frac{\lambda}{4} + \text{positive powers of } z.$$

$$-\frac{m^2}{z^2(z+2)^2} = -\frac{m^2}{z^2}\cdot\frac{1}{4\left(1+\frac{z}{2}\right)^2} = -\frac{m^2}{4z^2}\left(1+\frac{z}{2}\right)^{-2}$$

$$= -\frac{m^2}{4z^2}\left(1-2\frac{z}{2}+3\frac{z^2}{4}\right)$$

$$= -\frac{m^2}{4z^2} + \frac{m^2}{4z} - \frac{3m^2}{16} + \text{positive powers of } z.$$

Substituting these values in (4), we see that the lowest power of z is $\alpha - 2$ and that the coefficient of $z^{\alpha-2}$ is

$$\left\{\alpha(\alpha-1) + \alpha - \frac{m^2}{4}\right\} a_0,$$

$$\left(\alpha^2 - \frac{m^2}{4}\right) a_0. \tag{5}$$

Since a_0 can be taken to be different from zero, we must have :

$$\alpha^2 - \frac{m^2}{4} = 0. \tag{6}$$

This equation to determine the index α is called the "indicial equation". It has the roots

$$\alpha = +\frac{m}{2} \text{ and } \alpha = -\frac{m}{2}. \tag{7}$$

For the pole $x = -1$ also, we get the same indicial equation (6) and the same roots (7).

Since y has to be positive everywhere, α must be the positive root $\frac{m}{2}$.

For determining the coefficients a_p, it is convenient to change the dependent variable y by taking away from it the factor $(1-x^2)^{\frac{m}{2}}$. We write therefore

$$y = (1-x^2)^{\frac{m}{2}} v, \tag{8}$$

so that $x = -1$ and $x = +1$ are both ordinary points for v. Then

$$\frac{dy}{dx} = (1-x^2)^{\frac{m}{2}} \frac{dv}{dx} - mx\,(1-x^2)^{\frac{m}{2}-1} v,$$

$$\frac{d^2y}{dx^2} = (1-x^2)^{\frac{m}{2}} \frac{d^2v}{dx^2} - 2mx\,(1-x^2)^{\frac{m}{2}-1} \frac{dv}{dx} - m\,(1-x^2)^{\frac{m}{2}-1} v$$
$$+ m\,(m-2)\,x^2\,(1-x^2)^{\frac{m}{2}-2} v.$$

Substituting these values in (1) we get the equation

$$(1-x^2)^{\frac{m}{2}+1} \frac{d^2v}{dx^2} - 2\,(m+1)\,x\,(1-x^2)^{\frac{m}{2}} \frac{dv}{dx}$$
$$+ (\lambda - m - m^2)\,(1-x^2)^{\frac{m}{2}} v = 0,$$

or on dividing out by $(1-x^2)^{\frac{m}{2}}$,

$$(1-x^2)\frac{d^2v}{dx^2} - 2\,(m+1)\,x\frac{dv}{dx} + (\lambda - m - m^2)\,v = 0. \tag{9}$$

To integrate (9) we assume

$$v = \sum_{l=0}^{\infty} a_l x^l. \tag{10}$$

Substituting this in (9), and equating the coefficient of x^l to zero, we get the recurrence formula

$$(l+2)\,(l+1)\,a_{l+2} = \{l\,(l-1) + 2\,(m+1)\,l - \lambda + m + m^2\}\,a_l. \tag{11}$$

From (11) we see that the series (10) contains either only the even powers or only the odd powers of x.

Now we choose the parameter λ so that for $l = k$, the coefficient of a_l vanishes. Then all $a_{k+2}, a_{k+4}, \cdots$ will also vanish, and the series $\Sigma a_l x^l$ ends at $l = k$. Thus

$$k\,(k-1) + 2\,(m+1)\,k - \lambda + m + m^2 = 0,$$

or

$$\lambda = k(k-1) + 2(m+1)k + m + m^2$$

$$= (k+m)(k+m+1) \tag{12}$$

Let us write

$$n = k + m, \tag{13}$$

so that

$$\lambda = n(n+1). \tag{14}$$

Writing successively $k = 1, 2, 3, \cdots$, in (13), we get different discrete values of n, and then from (14) we get the whole set of the values of λ, giving the discrete set of the required eigen-values.

The essential feature of this method is terminating the series at some particular point, since the solution has to be finite for all values of x however large.

8 · 3. *Bechert's Method for Determining the Eigen-values.*

Another simple method for determining the eigen-values of differential equations has been given by Bechert. It has a great advantage over the polynomial method in that the eigen-values are obtained without actually solving the equation. It requires a study of the behaviour of the solution near singular points. We shall illustrate this method by taking a concrete example.

Consider the equation

$$\frac{d^2y}{dx^2} + (\lambda - x^2)y = 0, \tag{1}$$

where λ is the parameter. We shall see in the next section that the wave-equation of the linear harmonic oscillator can be reduced to this form.

Obviously, the only singular points of this equation are $x = -\infty$ and $x = +\infty$, all other points being ordinary points. We must therefore examine the behaviour of y in detail as x approaches $\pm\infty$.

Now when x is very large, λ can be neglected in comparison with x^2 and (1) reduces to

$$\frac{d^2y}{dx^2} - x^2y = 0. \tag{2}$$

It is easy to see that y behaves very nearly like $e^{-\frac{x^2}{2}}$ when x is large. For assuming

$$y = e^{-\frac{x^2}{2}} \tag{3}$$

we get

$$\frac{dy}{dx} = -\,x e^{-\frac{x^2}{2}},$$

$$\therefore \quad \frac{d^2y}{dx^2} = (x^2 - 1)\, e^{-\frac{x^2}{2}},$$

and since x is very large, we have

$$\frac{d^2y}{dx^2} = x^2 e^{-\frac{x^2}{2}} = x^2 y \text{ (very nearly)}.$$

Thus (3) satisfies very nearly the equation (2).

This shows that the asymptotic form of the solution y of (1) for $x \to \infty$ is $e^{-\frac{x^2}{2}}$. To solve the equation (1) exactly, we put therefore :

$$y = e^{-\frac{x^2}{2}} u. \tag{4}$$

Then

$$\frac{dy}{dx} = e^{-\frac{x^2}{2}} \left(\frac{du}{dx} - xu \right),$$

$$\frac{d^2y}{dx^2} = e^{-\frac{x^2}{2}} \left\{ \frac{d^2u}{dx^2} - 2x \frac{du}{dx} + (x^2 - 1)\, u \right\}.$$

Substituting these values of y and $\frac{d^2y}{dx^2}$ in (1), we get

$$e^{-\frac{x^2}{2}} \left\{ \frac{d^2u}{dx^2} - 2x \frac{du}{dx} + (x^2 - 1)\, u \right\} + (\lambda - x^2)\, e^{-\frac{x^2}{2}} u = 0,$$

or, simplifying and dividing out by $e^{-\frac{x^2}{2}}$,

$$\frac{d^2u}{dx^2} - 2x \frac{du}{dx} + (\lambda - 1)\, u = 0. \tag{5}$$

We have to study this equation when x is very large. We change the independent variable from x to z by writing

$$z = \frac{1}{x}, \tag{6}$$

so that we can study the behaviour when z is very small. Then

$$\frac{du}{dx} = -\,z^2 \frac{du}{dz}$$

$$\frac{d^2u}{dx^2} = z^4 \frac{d^2u}{dz^2} + 2z^3 \frac{du}{dz}.$$

On substituting these values in (5), it becomes

$$z^4 \frac{d^2u}{dz^2} + (2z^3 + 2z) \frac{du}{dz} + (\lambda - 1)\, u = 0. \tag{7}$$

Now when $x \rightarrow \infty$, $z \rightarrow 0$, and since z is very small, we can neglect z^3 and z^4 in comparison with z. The equation (7) then reduces to

$$2z \frac{du}{dz} + (\lambda - 1) u = 0. \tag{8}$$

This is a first order differential equation in which the variables are separable:

$$\frac{du}{u} + \frac{\lambda - 1}{z} \frac{dz}{z} = 0. \tag{9}$$

The integral of equation (9) is

$$\log u + \frac{\lambda - 1}{2} \log z = \log A$$

or

$$u z^{\frac{\lambda - 1}{2}} = A,$$

where A is an arbitrary constant. Writing again $z = \frac{1}{x}$, we get

$$u \left(\frac{1}{x}\right)^{\frac{\lambda - 1}{2}} = A$$

or

$$u = A x^{\frac{\lambda - 1}{2}}. \tag{10}$$

Since u must be finite for $x = 0$, the power of x, *viz.*, $\frac{\lambda - 1}{2}$ must be positive. Moreover, u must be single-valued, and this can be the case only when the power of x is a positive integer or zero. We have therefore:

$$\frac{\lambda - 1}{2} = n, \tag{11}$$

or

$$\lambda = (2n + 1), \quad (n = 0, 1, 2, 3, \cdots). \tag{12}$$

The eigen-values of the differential equations are 1, 3, 5, 7, · · ·, *i.e.*, the set of odd numbers.

8 · 4. *The Linear Harmonic Oscillator.*

We shall give a wave-mechanical treatment of the linear harmonic oscillator discussed in § 6 · 7 by the matrix mechanical method of Heisenberg.

If x is the displacement of the oscillator of mass μ from the position of equilibrium, then from (1) § 6 · 7 we know that the potential energy is

$$V = \tfrac{1}{2}kx^2,$$

where k is the elastic constant. From (4) of § 6 · 7 we have further

$$\frac{k}{\mu} = 4\pi^2\nu_0^2$$

where ν_0 is the frequency of the Oscillator. Therefore

$$V = 2\pi^2\nu_0^2\mu x^2. \tag{1}$$

For the linear oscillator, the problem is one-dimensional, and so the wave-equation is :

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2\mu}{h^2}(E - 2\pi^2\nu_0^2\mu x^2)\,\psi = 0. \tag{2}$$

We change the independent variable from x to q by taking

$$q = 2\pi\sqrt{\frac{\mu\nu_0}{h}}\,x. \tag{3}$$

The wave-equation (2) then becomes

$$\frac{4\pi^2\mu\nu_0}{h}\frac{d^2\psi}{dq^2} + \frac{8\pi^2\mu}{h}\left(E - 2\pi^2\nu_0^2\mu\cdot\frac{h}{4\pi^2\mu\nu_0}q^2\right)\psi = 0,$$

or, after simplifying,

$$\frac{d^2\psi}{dq^2} + \left(\frac{2E}{h\nu_0} - q^2\right)\psi = 0. \tag{4}$$

Now write

$$\lambda = \frac{2E}{h\nu_0}, \tag{5}$$

and the equation (4) reduces to

$$\frac{d^2\psi}{dq^2} + (\lambda - q^2)\,\psi = 0. \tag{6}$$

This is the same equation as (1) § 8 · 3, except that for the independent variable x we have q, and for the dependent variable y we have ψ.

From (12) of § 8 · 3, we know that the equation (6) has eigen-values given by

$$\lambda = (2n + 1). \tag{7}$$

Thus

$$\frac{2E}{h\nu_0} = 2n + 1.$$

The eigen-values are therefore

$$E_n = h\nu_0\,(n + \tfrac{1}{2}),\ (n = 0, 1, 2, \cdots), \tag{8}$$

giving again the half integral energy values required by experiments.

8 · 4 (1). *The Linear Harmonic Oscillator* (*Polynomial Method*).

To illustrate the polynomial method of finding eigen-values, we shall discuss the harmonic oscillator once more, and show that it leads to the same result.

The wave-equation of the oscillator is, from (2) § 8 · 4,

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2\mu}{h^2}(\mathrm{E} - 2\pi^2\nu_0{}^2\mu x^2)\,\psi = 0. \tag{1}$$

Let us write

$$\lambda = \frac{8\pi^2\mu}{h^2}\mathrm{E}, \quad a = \frac{4\pi^2\mu\nu_0}{h}; \tag{2}$$

then the equation (1) becomes

$$\frac{d^2\psi}{dx^2} + (\lambda - a^2x^2)\,\psi = 0. \tag{3}$$

The singular points of this equation are $x = -\infty$, and $x = +\infty$, and obviously, these are essentially singular points. To find the asymptotic behaviour of ψ for large x, we neglect λ in comparison with a^2x^2 in (3), and get the equation

$$\frac{d^2\psi}{dx^2} - a^2x^2\psi = 0. \tag{4}$$

It is easily seen that ψ behaves very much like $e^{-\frac{a}{2}x^2}$ for large x. For, writing

$$\psi = e^{-\frac{a}{2}x^2}, \tag{5}$$

we get

$$\frac{d\psi}{dx} = -ax\psi,$$

$$\frac{d^2\psi}{dx^2} = a^2x^2\,\psi - a\psi.$$

Since x is very large, we have

$$\frac{d^2\psi}{dx^2} = a^2x^2\psi \text{ (very nearly)},$$

and therefore (5) satisfies the equation (4). Thus the asymptotic form of the solution ψ of (3) for $x \rightarrow \infty$ is $e^{-\frac{a}{2}x^2}$. To solve the equation (3) exactly, we put

$$\psi = e^{-\frac{a}{2}x^2}u. \tag{6}$$

Then

$$\frac{d\psi}{dx} = e^{-\frac{\alpha}{2}x^2}\left(\frac{du}{dx} - \alpha xu\right),$$

$$\therefore \quad \frac{d^2\psi}{dx^2} = e^{-\frac{\alpha}{2}x^2}\left\{\frac{d^2u}{dx^2} - 2\alpha x\frac{du}{dx} - \alpha u + \alpha^2x^2u\right\}.$$

Substituting these values of ψ and $\dfrac{d^2\psi}{dx^2}$ in (3), we get:

$$e^{-\frac{\alpha}{2}x^2}\left(\frac{d^2u}{dx^2} - 2\alpha x\frac{du}{dx} - \alpha u + \alpha^2x^2u\right) + (\lambda - \alpha^2x^2)\, e^{-\frac{\alpha}{2}x^2}\, u = 0,$$

or, simplifying and dividing out by $e^{-\frac{\alpha}{2}x^2}$, we find

$$\frac{d^2u}{dx^2} - 2\alpha x\frac{du}{dx} + (\lambda - \alpha)\, u = 0. \tag{7}$$

We change the independent variable from x to z by writing

$$z = \sqrt{\alpha}\; x. \tag{8}$$

Then

$$\frac{du}{dx} = \sqrt{\alpha}\,\frac{du}{dz},\ \frac{d^2u}{dx^2} = \alpha\,\frac{d^2u}{dz^2}.$$

Therefore (7) becomes

$$\alpha\frac{d^2u}{dz^2} - 2\alpha z\frac{du}{dz} + \alpha\left(\frac{\lambda}{\alpha} - 1\right)u = 0,$$

or

$$\frac{d^2u}{dz^2} - 2z\frac{du}{dz} + \left(\frac{\lambda}{\alpha} - 1\right)u = 0. \tag{9}$$

We integrate this equation in series

$$u = \sum_{l=0}^{\infty} a_l\, z^l. \tag{10}$$

Substituting this in (9) and equating the coefficient of z^l to zero, we get the recurrence formula:

$$(l + 2)\,(l + 1)\, a_{l+2} + \left(\frac{\lambda}{\alpha} - 1 - 2l\right) a_l = 0. \tag{11}$$

Now, since z has to be finite even for large values, the series (10) must terminate at some point. Suppose the last non-vanishing term in the series (10) is a_nz^n. Then it is necessary that the factor of a_n in (11) must vanish. Therefore

$$\frac{\lambda}{\alpha} - 1 - 2n = 0,\quad (n = 0, 1, 2, \cdots),$$

or

$$\frac{\lambda}{\alpha} = 2n + 1, \quad (n = 0, 1, 2, 3, \cdots). \tag{12}$$

But from (2), we have

$$\frac{\lambda}{\alpha} = \frac{8\pi^2\mu E \times h}{h^2 \times 4\pi^2\mu\nu_0} = \frac{2E}{h\nu_0}. \tag{13}$$

Therefore

$$\frac{2E}{h\nu_0} = 2n + 1,$$

and the eigen-values are

$$E_n = h\nu_0 \left(n + \tfrac{1}{2}\right), \quad (n = 0, 1, 2, \cdots). \tag{14}$$

The corresponding eigen-functions are

$$\psi_n(x) = e^{-\frac{2\pi^2\mu\nu_0 x^2}{h}} H_n\left(2\pi\sqrt{\frac{\mu\nu_0}{h}}\,x\right), \tag{15}$$

where $H_n(z)$ is the "nth Hermitian Polynomial" given by the solution of the equation (7) in series :

$$u(z) = H_n(z) = (2z)^n - \frac{n(n-1)}{1\,!}(2z)^{n-2}$$

$$+ \frac{n(n-1)(n-2)(n-3)}{2\,!}(2z)^{n-4}$$

$$- \frac{n(n-1)(n-2)(n-3)(n-4)(n-5)}{3\,!}(2z)^{n-6} + \cdots \tag{16}$$

8 · 5. *The Hydrogen-like Atom.*

In his very first paper, Schrödinger treated the problem of the hydrogen atom by a wave-mechanical method. It was his success in deriving the energy values in agreement with those furnished by Bohr's theory that drew the attention of physicists to the wave-theory of matter enunciated previously by de Broglie.

We take a hydrogen-like atom, *i.e.*, a single electron of mass μ and charge $(-e)$ moving in the field of a nucleus of charge $+Ne$, where N is the atomic number of the element. The potential energy of the electron is therefore

$$V = -\frac{Ne^2}{r}, \qquad (1$$

where r is its distance from the nucleus. The wave-equation for the electron is then

$$\therefore \quad \frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}+\frac{8\pi^2\mu}{h^2}\left(E+\frac{Ne^2}{r}\right)\psi=0. \tag{2}$$

We transform to the spherical polar co-ordinates $r,\ \theta,\ \phi$ given by the relations

$$x=r\sin\theta\cos\phi,\ y=r\sin\theta\sin\phi,\ z=r\cos\theta\,; \tag{3}$$

then it is shown in the books on the theory of functions that

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)$$
$$+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}.$$

The wave-equation (2) then becomes in polar co-ordinates :

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)$$
$$+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}+\frac{8\pi^2\mu}{h^2}\left(E+\frac{Ne^2}{r}\right)\psi=0. \tag{4}$$

We assume that the solution of this differential equation is a product of three factors, each of which depends on a single-variable only. Thus let

$$\psi=R\,(r)\,f\,(\theta)\,F\,(\phi). \tag{5}$$

The conditions to be satisfied by ψ are that it mut be single valued, continuous and finite in the whole domain

$$0\leqslant r,\ 0\leqslant\theta\leqslant\pi,\ \ -\pi\leqslant\phi\leqslant\pi.$$

Now ϕ is a cyclic co-ordinate which does not enter explicitly in the equation (4). Therefore the function F (ϕ) is of the form

$$F\,(\phi)=e^{im\phi}, \tag{6}$$

where m is a constant which must be an integer because F (ϕ) has to be unique.

We write

$$g\,(r)=\frac{8\pi^2\mu}{h^2}\left(E+\frac{Ne^2}{r}\right). \tag{7}$$

Then substituting (5), (6), (7) in (4) and dividing out by R f F, we get the equation

$$\frac{r^2}{R}\left(\frac{d^2R}{dr^2}+\frac{2}{r}\frac{dR}{dr}+gR\right)+\frac{1}{f}\left\{\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{df}{d\theta}\right)-\frac{m^2f}{\sin^2\theta}\right\}=0,$$

or

$$\frac{r^2}{R}\left(\frac{d^2R}{dr^2}+\frac{2}{r}\frac{dR}{dr}+gR\right)=-\frac{1}{f}\left\{\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{df}{d\theta}\right)-\frac{m^2f}{\sin^2\theta}\right\}. \quad (8)$$

The left-hand side of (8) is quite independent of θ, whereas the right-hand side is quite independent of r, so that the two expressions must be independent of both r and θ. Thus, we must have:

$$\frac{r^2}{R}\left(\frac{d^2R}{dr^2}+\frac{2}{r}\frac{dR}{dr}+gR\right)=\lambda, \quad (9)$$

and

$$-\frac{1}{f}\left\{\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{df}{d\theta}\right)-\frac{m^2f}{\sin^2\theta}\right\}=\lambda, \quad (10)$$

where λ is an absolute constant.

Let us consider the equation (10) first. It is

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{df}{d\theta}\right)+\left(\lambda-\frac{m^2}{\sin^2\theta}\right)f=0. \quad (11)$$

Transform the variables θ, f into x, y by writing

$$x=\cos\theta,\quad y(x)=f(\theta), \quad (12)$$

then

$$\frac{df}{d\theta}=\frac{df}{dx}\cdot\frac{dx}{d\theta}=\frac{df}{dx}(-\sin\theta);$$

$$\therefore\quad \frac{1}{\sin\theta}\frac{df}{d\theta}=-\frac{df}{dx};\quad \sin\theta\frac{df}{d\theta}=-\sin^2\theta\frac{df}{dx}=-(1-x^2)\frac{df}{dx}.$$

Substituting in (11), we get

$$-\frac{d}{dx}\left\{-(1-x^2)\frac{dy}{dx}\right\}+\left\{\lambda-\frac{m^2}{1-x^2}\right\}y=0,$$

or

$$(1-x^2)\frac{d^2y}{dx^2}-2x\frac{dy}{dx}+\left(\lambda-\frac{m^2}{1-x^2}\right)y=0. \quad (13)$$

This equation (13) is the same as (1) § 8 · 2. We found in that section, that the eigen-values of this equation are

$$\lambda=l\,(l+1), \quad (14)$$

where λ is a positive integer not less than m.

The eigen-functions are given by (8), (10) § 8 · 2:

$$y=(1-x^2)^{\frac{m}{2}}v;$$

where $x=\cos\theta$, and

$$v=\Sigma\, a_k\, x^k.$$

The coefficients a_k are determined by the recurrence formula (11) § 8 · 2. For a particular value of l, this solution comes out to be

$$f(\theta) = y = \frac{(2l)!\sin^m\theta}{2(l-m)!}\left\{\cos^{l-m}\theta - \frac{(l-m)(l-m-1)}{2(2l-1)}\cos^{l-m-2}\theta\right.$$
$$\left.+\frac{(l-m)(l-m-1)(l-m-2)(l-m-3)}{2\cdot 4\cdot(2l-1)(2l-3)}\cos^{l-m-4}\theta + \cdots\right\} \quad (15)$$

This solution (15) is denoted by P_l^m (cos θ), and is called the "associated Legendre function".

Thus the two factors $f(\theta)$, $F(\phi)$ in the solution (5) become on account of (6) and (15) :

$$S_l(\theta, \phi) = f(\theta)F(\phi) = e^{im\phi}P_l^m(\cos\theta). \quad (16)$$

The function $S_l(\theta, \phi)$ is called the "Surface Spherical Harmonics".

We have now to find that factor of the wave-function ψ which depends on r. For this purpose we have the equation (9). This is :

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left(g - \frac{\lambda}{r^2}\right)R = 0,$$

or substituting the values of g from (7) :

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left(\frac{8\pi^2\mu}{h^2}E + \frac{8\pi^2\mu Ne^2}{h^2}\frac{1}{r} - \frac{\lambda}{r^2}\right)R = 0.$$

On writing

$$A = \frac{8\pi^2\mu}{h^2}E, \quad B = \frac{4\pi^2\mu Ne^2}{h^2}, \quad C = -\lambda = -l(l+1), \quad (17)$$

the equation becomes

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left(A + 2\frac{B}{r} + \frac{C}{r^2}\right)R = 0. \quad (18)$$

If we wish to find the discrete spectrum of the hydrogen-like atom, we must take E negative. Then A is also negative, and we can write

$$A = -\frac{1}{r_0^2}. \quad (19)$$

In order to determine the asymptotic form of the solution for large r, we neglect terms with $\frac{1}{r}$ and $\frac{1}{r^2}$ in (18) and get

$$\frac{d^2R}{dr^2} + AR = \frac{d^2R}{dr^2} - \frac{R}{r_0^2} = 0. \quad (20)$$

The solutions of (20) are of the form

$$R = e^{\pm \frac{r}{r_0}};$$

but since r must be finite when $r \rightarrow \infty$, we must take

$$R = e^{-\frac{r}{r_0}}. \tag{21}$$

We change the independent variable from r to z by writing

$$z = 2\frac{r}{r_0} = 2\sqrt{-A}\,r, \quad (0 \leqslant z < \infty) \tag{22}$$

The asymptotic solution (21) then becomes :

$$R = e^{-\frac{z}{2}}. \tag{23}$$

Further the equation (18) is transformed to

$$\frac{4}{r_0^2}\frac{d^2R}{dz^2} + \frac{4}{zr_0}\cdot\frac{2}{r_0}\frac{dR}{dz} + \left\{-\frac{1}{r_0^2} + \frac{4B}{zr_0} - \frac{4l(l+1)}{z^2r_0^2}\right\}R = 0,$$

or, dividing out by $\frac{4}{r_0^2}$,

$$\frac{d^2R}{dz^2} + \frac{2}{z}\frac{dR}{dz} + \left\{-\tfrac{1}{4} + \frac{B}{r_0}\tfrac{1}{2} - \frac{l(l+1)}{z^2}\right\}R = 0. \tag{24}$$

For the solution of this equation, we set

$$R = e^{-\frac{z}{2}}v(z), \tag{25}$$

then

$$\frac{dR}{dz} = e^{-\frac{z}{2}}\left(\frac{dv}{dz} - \tfrac{1}{2}\,v\right),$$

$$\frac{d^2R}{dz^2} = e^{-\frac{z}{2}}\left(\frac{d^2v}{dz^2} - \frac{dv}{dz} + \tfrac{1}{4}\,v\right).$$

Substituting these in (24) and simplifying, we have :

$$\frac{d^2v}{dz^2} + \left(\frac{2}{z} - 1\right)\frac{dv}{dz} + \left\{\left(\frac{B}{r_0} - 1\right)\frac{1}{z} - \frac{l(l+1)}{z^2}\right\}v = 0. \tag{26}$$

The only singular point of this equation is $z = 0$, and it is a pole. Therefore, to solve this equation, we assume the series

$$v = z^p \sum_{k=0}^{\infty} a_k z^k, \tag{27}$$

where p is a number which has to be determined. If we substitute (26) in (27), we find that the term with the lowest power of z is

$$\{p\,(p-1) + 2p - l\,(l+1)\}\,a_0,$$

or

$$\{p\,(p+1) - l\,(l+1)\}\, a_0.$$

Since a_0 can be assumed to be different from zero, we get the indicial equation :

$$p\,(p+1) - l\,(l+1) = 0. \tag{28}$$

The solutions of this equation are

$$\text{either } p = l \quad \text{or} \quad p = -(l+1).$$

Since for $z = 0$, v must be finite, we must take the first solution only :

$$p = l. \tag{29}$$

The solution of the equation (26) can therefore be written

$$v = z^l \sum_{k=0}^{\infty} a_k z^k. \tag{30}$$

Substituting (30) in (26) and equating the coefficient of z^{k+l-1} to zero, we get the recurrence formula for determining the a_k :

$$\{(k+l+1)\,(k+l) + 2\,(k+l+1) - l\,(l+1)\}\, a_{k+1}$$
$$= \left(k + l + 1 - \frac{B}{r_0}\right) a_k. \tag{31}$$

In order that v remains finite even for large z, it is necessary that v should be a polynomial, *i.e.*, the series (30) must terminate at some point. This will be the case only when in (31), the coefficient of a_k vanishes for some value n_r of k. Thus

$$\frac{B}{r_0} = n_r + l + 1, \quad (n_r = 0, 1, 2, \cdots). \tag{32}$$

We set

$$n = n_r + l + 1, \tag{33}$$

and see that n can have any of the values $1, 2, \cdots$, *i.e.*, n can be any positive integer. Then from (32) we get on squaring

$$\frac{B^2}{r_0^2} = n^2,$$

or, substituting the values from (17) and (19) :

$$n^2 = \frac{B^2}{r_0^2} = -\frac{B^2}{A} = -\frac{16\pi^4\mu^2N^2e^4}{h^4} \times \frac{h^2}{8\pi^2\mu E} = -\frac{2\pi^2\mu N^2e^4}{h^2E}.$$

Consequently,

$$E = -\frac{2\pi^2\mu N^2e^4}{h^2n^2} \quad (n = 1, 2, 3, \cdots). \tag{34}$$

These are the discrete energy values, and we see that they are identical with the values (15) § 5 · 3 given by Bohr. They agree, therefore, with experiment.

If we take E positive, we find that it remains inditerminate. The theory therefore leads in this case to a continuous spectrum, again in agreement with experiment.

Heisenberg's Matrix Mechanics also leads to a discrete as well as a continuous spectrum.

As remarked by Sommerfeld, "it is a particularly beautiful feature of Schrödinger's theory, that it connects the continuous spectrum by a uniform analytical process with the line spectrum of hydrogen. Bohr's quantum theory required special assumptions to achieve this result".

8 · 6. *The Equivalence of Matrix Mechanics and Wave-Mechanics.*

We have seen that the energy values of the linear harmonic oscillator come out to be the same when treated both by the Matrix method and the Wave method. It was found that for other problems also the results arrived at were essentially the same. Now the two methods have apparently quite opposite characters. The matrix theory is " a true theory of a discontinuum ", in which the classical continuous variables are replaced by systems of discrete numerical quantities which are defined by algebraic equations. On the other hand, wave-mechanics replaces the classical point-mechanics by a continuum theory, where we have a continuous field-like process governed by a single partial-differential equation. In view of this apparent contradiction in the methods of the two theories, it is surprising that both of them should lead to the same results. It was Schrödinger himself who solved this puzzle by showing that the two methods were mathematically equivalent. In fact, he proved that we can deduce the Heisenberg matrices from the wave-functions, and that conversely, we can determine the wave-functions from the Heisenberg matrices.

This discovery was of great importance in the subsequent development of the quantum mechanics. It furnished a method for calculating the matrix elements by simple integration.

We shall prove this equivalence in the following two subsections :

8 · 6 (1). *The Wave-Equation.*

Let us first take a dynamical system having one degree of freedom, and suppose that q is the co-ordinate and p the conjugate momentum. Then the quantum condition of Heisenberg's matrix mechanics is

$$pq - qp = \frac{h}{2\pi i}\,1, \tag{1}$$

where 1 is the unit matrix.

Now let $f\,(q)$ be any function of q. Then

$$\frac{d}{dq}\,(qf) = f + q\,\frac{df}{dq},$$

or, rearranging, and taking f outside :

$$\left(\frac{d}{dq}\,q - q\,\frac{d}{dq}\right) f = f.$$

This means that

$$\frac{d}{dq}\,q - q\,\frac{d}{dq} = 1, \tag{2}$$

showing that $\frac{d}{dq}\,q - q\,\frac{d}{dq}$ is a unit operator. Multiplying (1) by the constant $\frac{h}{2\pi i}$, we get

$$\frac{h}{2\pi i}\,\frac{d}{dq}\cdot q - q\cdot\frac{h}{2\pi i}\,\frac{d}{dq} = \frac{h}{2\pi i}\cdot 1. \tag{3}$$

Now, a comparison of (1) with (3) suggests that the momentum p can be taken as a sort of differential operator

$$p = \frac{h}{2\pi i}\,\frac{d}{dq}. \tag{4}$$

In general, if we have a particle of mass m with co-ordinates $(x,\ y,\ z)$ and velocities $(u,\ v,\ w)$, then we must have according to (4),

$$\left.\begin{aligned} mu &= m\dot{x} = \frac{h}{2\pi i}\,\frac{\partial}{\partial x}, \\ mv &= m\dot{y} = \frac{h}{2\pi i}\,\frac{\partial}{\partial y}, \\ mw &= m\dot{z} = \frac{h}{2\pi i}\,\frac{\partial}{\partial z}. \end{aligned}\right\} \tag{5}$$

But according to the relativity theory, the time t has the same standing, and any picture of nature must be symmetrical in the four co-ordinates x, y, z, ict. From that theory we know also that if the energy is E, then $-E$ is the momentum conjugate to the co-ordinate t. This is because p_x, p_y, p_z, E are the components of a four-dimensional vector, where $E = mc^2$. Thus we must add a fourth equation to the three equations (5), given by

$$-E = \frac{h}{2\pi i}\frac{\partial}{\partial t}$$

or

$$mc^2 = E = -\frac{h}{2\pi i}\frac{\partial}{\partial t}. \tag{6}$$

We consider again a system having only one degree of freedom characterised by the co-ordinate x. We perform the operations on a function ψ. From the first equation of (5) we have

$$mu\psi = \frac{h}{2\pi i}\frac{\partial\psi}{\partial x},$$

or

$$\frac{\partial\psi}{\partial x} = 2\pi i \cdot \frac{mu}{h}\psi, \tag{7}$$

where, for the present, we take the velocity u to be uniform. The solution of (7) is

$$\psi = \psi_0 e^{2\pi i \frac{mu}{h} x}, \tag{8}$$

giving ψ as a wave of wave-length $\lambda = \frac{h}{mu}$. This shows that the momentum mu of the moving particle has associated with it a train of regular waves of wave-length $\frac{h}{mu}$.

Now apply the operator (6) to the function ψ:

$$E\psi = -\frac{h}{2\pi i}\frac{\partial\psi}{\partial t},$$

or

$$\frac{\partial\psi}{\partial t} = -2\pi i\frac{E}{h}\psi = -2\pi i\frac{mc^2}{h}\psi, \tag{9}$$

giving for the solution :

$$\psi = \psi_0 e^{-2\pi i \frac{E}{h} t} = \psi_0 e^{-2\pi i \frac{mc^2}{h} t}. \tag{10}$$

The waves have vibrations in time of period $\frac{h}{E} = \frac{h}{mc^2}$, *i.e.*, of frequency $\nu = \frac{E}{h} = \frac{mc^2}{h}\cdot$ Thus the material particle has associated with it vibrations in time of frequency $\nu = \frac{E}{h} = \frac{mc^2}{h}\cdot$ Combining the two formulas (8) and (10) we have

$$\psi = \psi_0 \, e^{2\pi i \frac{mu}{h}} \left(x - \frac{c^2}{u} t\right). \tag{11}$$

This represents waves of wave-length $\lambda = \frac{mu}{h}$ travelling in the direction of x with a speed $\frac{c^2}{u}\cdot$ The phase velocity of the system of waves is therefore $\frac{c^2}{u}\cdot$ We can deduce in this way all the results of de Broglie's original theory.

If the particle is moving in three-dimensions (x, y, z) with a variable velocity (u, v, w), in a conservative field of force having a potential $V(x, y, z)$, then, if T is the kinetic energy:

$$T = \tfrac{1}{2}m\,(u^2 + v^2 + w^2) = E - V. \tag{12}$$

Substituting in this the three equations (5), we get

$$-\frac{h^2}{8\pi^2 m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) = E - V. \tag{13}$$

Performing the operation (13) on a function ψ, and rearranging, we get

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{8\pi^2 m}{h^2}(E - V)\,\psi = 0. \tag{14}$$

Thus we get again Schrödinger's equation, showing that a material particle moving in a conservative field has associated with it a train of waves whose wave-length, however, varies with V, just as the wave-length of light varies as it passes through a refracting substance.

8 · 6 (2). *Derivation of the Matrix Elements from the Eigen-functions.*

For a particle of mass m_0, moving in a field of potential V, we had in (12) § 8 · 6 (1):

$$\tfrac{1}{2}m_0\,(u^2 + v^2 + w^2) = E - V,$$

where E is the total energy. Writing $T = \tfrac{1}{2}m_0\,(u^2 + v^2 + w^2)$, and rearranging, we can put this equation in the form

$$T + V = E.$$

Let q_r $(r = 1, 2, \cdots, s)$ be the generalised co-ordinates of the dynamical system, and p_r the corresponding conjugate momenta. If we express T in q_r, p_r, and V in q_r, the above equation of energy becomes

$$H(q_r, p_r) - E = 0, \tag{1}$$

where $H(q_r, p_r)$ is the Hamiltonian of the dynamical system. Now to make a transition from the particle mechanics to wave-mechanics, we found that it is necessary to write

$$p_r = \frac{h}{2\pi i}\frac{\partial}{\partial q_r}, \quad (r = 1, 2, \cdots, s), \tag{2}$$

$$E = -\frac{h}{2\pi i}\frac{\partial}{\partial t}. \tag{3}$$

Thus both H and E are differential operators. If we perform these operations on the wave-function ψ, the wave-equation becomes :

$$(H - E)\,\psi = 0, \tag{4}$$

where H denotes the operator

$$H\left(q_r, \frac{h}{2\pi i}\frac{\partial}{\partial q_r}\right), \tag{5}$$

and E denotes the operator (3).

From the general solution of the differential equation (4), the matrices that form the solution of the dynamical problem may be very easily obtained.

Let

$$\psi_n, \quad (n = 1, 2, \cdots) \tag{6}$$

be a complete independent set of orthogonal and normalised eigen-functions of the differential equation (4). Each of these terms has a definite meaning. Thus by saying that the set ψ_n is orthogonal, we mean that for any m and any n :

$$\int \psi_m^* \psi_n dq = 0, \quad (m \neq n), \tag{7}$$

where the integral extends over the whole q space, and where ψ_m^* denotes the conjugate of ψ_m.

Similarly, the property of being normal requires that

$$\int \psi_n^* \psi_n dq = \int |\psi_n|^2 dq = 1, \tag{8}$$
$$(n = 1, 2, 3, \cdots).$$

The properties (7) and (8) are analogous to the properties possessed by the set of trigonometric functions:

$$\sin n\pi\nu t, \cos n\pi\nu t \quad (n = 1, 2, 3, \cdots).$$

Finally, just as every function which is periodic, and which satisfies certain conditions as regards continuity, can be expressed as a Fourier series, similarly, every function which satisfies the equation (4) can be expressed as a series of the eigen-functions ψ_n. This property is expressed by saying that the set of eigen-functions ψ_n "is complete," and it holds because of the linear character of the equation (4).

Thus the general solution ψ of equation (4) will be of the form

$$\psi = \sum_n c_n \psi_n, \tag{9}$$

where the c_n are the "Fourier coefficients" of the expansion, and are determined by the same method used for the Fourier series. Thus if we wish to determine c_m, we multiply the equation (9) by ψ_m^* and integrate both sides over the whole of q space, giving

$$\int \psi_m^* \psi dq = \sum_n c_n \int \psi_m^* \psi_n dq.$$

In the sum on the right-hand sides, all the integrals for which $n \neq m$ vanish on account of (7), so that

$$\int \psi_m^* \psi dq = c_m \int \psi_m^* \psi_m dq.$$

On account of (8), this gives for all $m \geqslant 1$

$$c_m = \int \psi_m^* \psi dq. \tag{10}$$

We shall prove now that any constant of integration of the dynamical system can be represented by a matrix whose elements are constants, there being one row and column of the matrix corresponding to each eigen-function ψ_n.

If a is any constant of integration of the system, then from (5) § 6 · 5,

$$\dot{a} = [a, \mathrm{H}],$$

where $\dot{a}$ is the time derivative of the matrix a, and $[a, \mathrm{H}]$ is the Poisson Bracket. Therefore, since a is constant in time, $[a, \mathrm{H}] = 0$, so that a must commute with H, giving $a\mathrm{H} = \mathrm{H}a$. Similarly, since E is the total energy, a commutes also with E, *i.e.*, $a\mathrm{E} = \mathrm{E}a$. Thus we have the relation

$$[a, \mathrm{H} - \mathrm{E}] = 0,$$

or

$$a\,(\mathrm{H} - \mathrm{E}) = (\mathrm{H} - \mathrm{E})\,a. \tag{11}$$

If we apply both the operators in (11) on the eigen-function ψ_n, we get for all $n \geqslant 1$:

$$a\,(\mathrm{H} - \mathrm{E})\,\psi_n = (\mathrm{H} - \mathrm{E})\,a\,\psi_n. \tag{12}$$

But, since ψ_n is a solution of the differential equation (4), we have $(\mathrm{H} - \mathrm{E})\,\psi_n = 0$, so that for all n

$$(\mathrm{H} - \mathrm{E})\,a\,\psi_n = 0, \tag{13}$$

i.e., $a\,\psi_n$ is also a solution of (4).

But we have seen in (9) that every solution of (4) can be expressed as a series in ψ_n of the type (9), so that

$$a\psi_n = \sum_m a_{mn}\psi_m, \tag{14}$$

where the a_{mn} are the constant development coefficients, which, according to (10), are given by the integral

$$a_{mn} = \int \psi_m{}^* a\psi_n dq. \tag{15}$$

We shall prove now that the quantities a_{mn} are the elements of the Heisenberg matrix belonging to a.

First of all, we see that

$$a_{nm} = \int \psi_n{}^* a\psi_m\, dq,$$

so that if a is a real quantity, then for all m, n:

$$\begin{aligned} a_{nm}{}^* &= \int (\psi_n{}^* a\psi_m)^*\, dq, \\ &= \int \psi_m{}^* a\psi_n\, dq \\ &= a_{mn}, \end{aligned}$$

showing that the matrix a is Hermitian.

Now, suppose b is another constant of integration, then as in (14), we get

$$b\,\psi_n = \sum_m b_{mn}\,\psi_m, \quad (n = 1, 2, 3, \cdots), \tag{16}$$

where

$$b_{mn} = \int \psi_m{}^* b\psi_n\, dq. \tag{17}$$

If a_{mn} and b_{mn} are the elements of Heisenberg matrices, then the laws of matrix addition and multiplication must hold. Now

$$\begin{aligned} (a + b)_{mn} &= \int \psi_m{}^*\,(a + b)\,\psi_n\, dq, \\ &= \int \psi_m{}^* a\psi_n\, dq + \int \psi_m{}^* b\psi_n\, dq \\ &= a_{mn} + b_{mn}\,; \end{aligned}$$

the law of matrix addition holds.

Similarly,

$$(ab)_{mn} = \int \psi_m{}^* ab\psi_n \, dq.$$

But from (14), by taking the conjugate series :

$$\psi_m{}^* a = \sum_k \psi_k{}^* a_{mk}$$

and from (17)

$$b\psi_n = \sum_j b_{jn}\psi_j.$$

Therefore

$$(ab)_{mn} = \int \sum_k \sum_j \psi_k{}^* a_{mk} b_{jn} \psi_j \, dq.$$

But, on account of the orthogonal and normal properties (7) and (8), this integral is only then different from zero for each k when $j = k$, so that

$$(ab)_{mn} = \sum_k a_{mk} b_{kn},$$

which is the rule for matrix multiplication.

Thus we can show that as soon as we have solved the Schrödinger equation completely, determining all the eigen-functions, we have found the solution of the quantum mechanical problem also. For, we can determine the matrices $q = (q_{mn})$ and $p = (p_{mn})$ of the quantum mechanical problem according to

$$q_{mn} = \int \psi_m{}^* q\psi_n \, dq, \tag{18}$$

$$p_{mn} = \int \psi_m{}^* p\psi_n \, dq = \int \psi_m{}^* \frac{h}{2\pi i} \frac{\partial}{\partial q} \psi_n \, dq. \tag{19}$$

To establish the equivalence of the two methods, we have to prove, of course, that if the canonical conjugate variables q, p are defined according to (18) and (19), then the following four relations hold :

(*a*) the quantum condition

$$pq - qp = \frac{h}{2\pi i} 1 ; \tag{20}$$

(*b*) any function V (q) of q can be defined as :

$$V_{mn}(q) = \int \psi_m{}^* V\psi_n \, dq ; \tag{21}$$

(*c*) the function p^2 is given by :

$$(p^2)_{mn} = \int \psi_m{}^* p^2 \psi_n \, dq = \int \psi_m{}^* \left(\frac{h}{2\pi i} \frac{\partial}{\partial q}\right)^2 \psi_n \, dq ; \tag{22}$$

(*e*) and finally that the Hamiltonian H (p, q) :

$$H(p, q) = \frac{1}{2m_0} \sum_1^s p^2 + V(q), \tag{23}$$

is a diagonal matrix. Here the function V (q) is taken as the potential function, and s is the degree of freedom.

We shall prove these relations one by one.

(*a*) From the law of multiplication for matrices, we have:

$$(pq)_{mn} = \sum_k p_{mk}q_{kn} = \frac{h}{2\pi i}\int \psi_m^* \frac{\partial}{\partial q} \sum_k \psi_k q_{kn} \cdot dq,$$

where we have written for p_{mk} its value from (19). But from (14), on writing q for a and k for m, we see that $q\psi_n = \sum_k q_{kn}\psi_k$, so that the last equation becomes:

$$(pq)_{mn} = \frac{h}{2\pi i}\int \psi_m^* \frac{\partial}{\partial q} q\psi_n \cdot dq. \tag{24}$$

Similarly, we have

$$\begin{aligned}(qp)_{mn} &= \sum_k q_{mk}p_{kn} = \int \psi_m^* q \sum_k \psi_k p_{kn} \cdot dq \\ &= \int \psi_m^* q p \psi_n \cdot dq = \frac{h}{2\pi i}\int \psi_m^* q \frac{\partial}{\partial q}\psi_n \cdot dq. \end{aligned} \tag{25}$$

Thus from (24) and (25):

$$(pq - qp)_{mn} = \frac{h}{2\pi i}\int \psi_m^* \left\{\frac{\partial}{\partial q} q - q\frac{\partial}{\partial q}\right\} \psi_n \cdot dq. \tag{26}$$

In (2) § 8·6 (1) we have shown that $\frac{\partial}{\partial q} q - q\frac{\partial}{\partial q} = 1$, so that

$$\begin{aligned}(pq - qp)_{mn} &= \frac{h}{2\pi i}\int \psi_m^* \psi_n \, dq \\ &= \frac{h}{2\pi i}\,\delta_{mn}, \end{aligned} \tag{27}$$

on account of (7) and (8). This proves (20).

(*b*) We show that if (21) is true for the function V $(q) = q^r$, it will be true for q^{r+1} also. For,

$$\begin{aligned}(q^{r+1})_{mn} &= (q^r \cdot q)_{mn} = \sum_k (q^r)_{mk}q_{kn} \\ &= \int \psi_m^* q^r \sum_k \psi_k q_{kn}\, dq = \int \psi_m^* q^r \cdot q\psi_n \cdot dq. \\ &= \int \psi_m^* q^{r+1}\psi_n \, dq. \end{aligned} \tag{28}$$

Using Induction, we see that (21) holds for all powers of q. Now since any analytical function V (q) can be expressed as a power series V $(q) = \sum c_r q^r$, we conclude that (21) holds for any such function.

(*c*) We have

$$
\begin{aligned}
(p^2)_{mn} &= \sum_k p_{mk}p_{kn} = \int \psi_m^* \frac{h}{2\pi i}\frac{\partial}{\partial q}\sum_k \psi_k p_{kn}\cdot dq \\
&= \int \psi_m^* \frac{h}{2\pi i}\frac{\partial}{\partial q} p\psi_n\cdot dq = \int \psi_m^* \frac{h}{2\pi i}\frac{\partial}{\partial q}\frac{h}{2\pi i}\frac{\partial}{\partial q}\psi_n\, dq \\
&= \int \psi_m^* \left(\frac{h}{2\pi i}\right)^2 \frac{\partial^2}{\partial q^2}\psi_n\, dq = \int \psi_m^* \frac{-h^2}{4\pi^2}\frac{\partial^2}{\partial q^2}\psi_n\, dq,
\end{aligned}
$$

which establishes (22).

(*d*) We are in a position now to prove that the Hamiltonian H given by (23) is a diagonal matrix. For, we have

$$
\begin{aligned}
\mathrm{H}_{mn} &= \int \psi_m^*\, \mathrm{H}\psi_n dq = \int \psi_m^* \left\{\frac{1}{2m_0}\sum_1^s p^2 + \mathrm{V}(q)\right\}\psi_n\, dq \\
&= \int \psi_m^* \left\{\frac{1}{2m_0}\sum_1^s\left(\frac{-h^2}{4\pi^2}\frac{\partial^2}{\partial q^2}\right) + \mathrm{V}(q)\right\}\psi_n\cdot dq. \qquad (29)
\end{aligned}
$$

Now, for a particle of mass m_0 moving in a field of potential V (q), the eigen-values E_n and the eigen-functions ψ_n are given by the Schrödinger equation

$$
\left\{\frac{1}{2m_0}\sum_1^s\left(\frac{-h^2}{4\pi^2}\frac{\partial^2}{\partial q^2}\right) + \mathrm{V}(q)\right\}\psi_n = \mathrm{E}_n\psi_n,
$$

so that (29) becomes :

$$
\begin{aligned}
\mathrm{H}_{mn} &= \int \psi_m^*\mathrm{E}_n\psi_n\, dq = \mathrm{E}_n \int \psi_m^*\psi_n\, dq \\
&= \mathrm{E}_n\, \delta_{mn}. \qquad (30)
\end{aligned}
$$

Thus we see that H is a diagonal matrix whose diagonal elements are $\mathrm{E}_1, \mathrm{E}_2, \cdots$, *i.e.*, exactly the same as the eigen-values supplied by the wave-mechanics.

The two methods are therefore completely equivalent.

CHAPTER IX.

PERTURBATION THEORY.

9·1. *Outline of the Perturbation Method.*

IN the preceding chapter, we have given a method of determining the eigen-values and eigen-functions of differential equations, and applied it to one or two simple physical problems. But, even for slightly more complicated problems, such as the problem of the hydrogen atom in an external electric field, this method of determining the eigen-values and eigen-functions is not applicable. It becomes necessary in these cases to develop approximate methods of finding the eigen-values and eigen-functions.

The first and most important approximate method in wave-mechanics was given by Schrödinger, though a similar method in matrix mechanics was developed by Born, Heisenberg and Jordan previously. This is a perturbation method having close resemblance to the method used for the problem of three bodies in astronomy, and described in detail by H. Poincaré in his classical treatise ***Méthodes nouvelles de la Mécanique Céleste.*** In 1918 Bohr adopted this method for his quantum theory, and his school applied it successfully to the relativistic stark-effect and other problems. But the theory of perturbations developed by Schrödinger is much simpler and clearer than either of these.

The object of the perturbation method is to increase the range of application of the eigen-value theory to problems which are not directly soluble by the comparatively elementary methods of the preceding chapter, but which are closely related to a directly soluble problem.

In its main outline, the method consists in breaking up the Hamiltonian H of the given problem into two parts, H_0 and ϵH_1, of which H_0 is simple and ϵH_1 is small. The first part H_0 can be considered as the Hamiltonian of a simple problem of the kind discussed in the preceding chapter, whose eigen-values and eigen-functions can be determined exactly. The addition of the small term ϵH_1 will then require small corrections, in the nature of a perturbation, in the eigen-values and eigen-functions of the unperturbed H_0 problem. ϵ is a small numerical factor, and the solution of the equation for the perturbed system can be obtained in the form of a power series in ϵ, which may be broken up at any point to get the desired degree of approximation, provided the series converges. Even if it does not converge, the first approximation has been found usually to give fairly accurate results.

The success of the perturbation method is due to the fact, that the eigen-values and eigen-functions vary continuously with the coefficients of the differential equations. This is an important theorem in the theory of boundary value problems.

In order to illustrate the various points in the theory, we shall first discuss the example of the vibrating membrane.

9·2. *The Vibrating Membrane.*

We consider a rectangular membrane whose edges are fixed in the straight lines $x = 0$, $x = a$, $y = 0$, $y = b$, in the (x, y) plane. The differential equation for the homogeneous membrane is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial t^2}. \tag{1}$$

For the solution of this equation, we write

$$u\,(x, y, t)) = v\,(x, y)\,g\,(t), \tag{2}$$

and substituting (2) in (1), we get

$$g\,(t)\left\{\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right\} = v\,(x, y)\,\frac{d^2 g}{dt^2},$$

or dividing both sides by $v\,(x, y)\,g\,(t)$:

$$\frac{1}{v\,(x, y)}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) = \frac{1}{g\,(t)}\,\frac{d^2 g}{dt^2}.$$

The right-hand side of this equation does not depend on x, y, and the left-hand side does not depend on t, so that each of them must be independent of both x, y and of t. Thus

$$\frac{1}{g\,(t)}\,\frac{d^2 g}{dt^2} = -\lambda, \tag{3}$$

$$\frac{1}{v\,(x, y)}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) = -\lambda, \tag{4}$$

where λ is an absolute constant which must be positive.

The equation (3) then gives

$$\frac{d^2 g}{dt^2} + \lambda g = 0,$$

whose solution is

$$g(t) = \mathrm{A} \cos \sqrt{\lambda}\, t + \mathrm{B} \sin \sqrt{\lambda}\, t. \tag{5}$$

Equation (4) now becomes

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \lambda v = 0, \tag{6}$$

with the boundary conditions

$$\begin{aligned} v &= 0 \text{ for } x = 0 \text{ and } x = a, \text{ for all } y, \\ v &= 0 \text{ for } y = 0 \text{ and } y = b, \text{ for all } x, \end{aligned} \tag{7}$$

while λ is the eigen-value parameter.

It is easily verified that the solutions, *i.e.*, eigen-functions, of (6) are,

$$v_{m,n}(x, y) = \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b}, \tag{8}$$

and the eigen-values

$$\lambda_{m,n} = \pi^2 \left(\frac{m^2}{a^2} + \frac{n^2}{b^2} \right), \tag{9}$$

$$(m, n = 1, 2, 3, \cdots).$$

If a and b are incommensurable, there is no other pair m', n' such that $\lambda_{m', n'} = \lambda_{m, n}$. To each eigen-value $\lambda_{m, n}$ then corresponds only one eigen-function $v_{m,n}(x, y)$. Each eigen-value is called "simple", and the case is non-degenerate.

Now, suppose that an external transversal pressure $\mathrm{P}(x, y)$ acts on the membrane, distributed anyhow over it; then the equation (6) becomes

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \lambda v = \mathrm{P}(x, y), \tag{10}$$

giving a non-homogeneous equation.

If the pressure P is periodic, with the period ω, and if ω coincides with one of the eigen-vibrations of the membrane, then it will excite the membrane to continually increasing vibrations, so that the disturbance will finally grow beyond all limits. If this "Catastrophe" is not to occur, there must be no resonance, and therefore P cannot on the whole perform any work on the vibrating membrane. To represent the vanishing of this work mathematically, we notice that if $d\mathrm{S}$ is an element of the membrane acted upon by the pressure P, then since the displacement from the position of rest is v, the condition is

$$\iint \mathrm{P}v \, d\mathrm{S} = 0, \tag{11}$$

i.e., that the right-hand side of equation (10) must be orthogonal to the eigen-function of the left-hand side equated to zero. This theorem holds in general.

Suppose now that the membrane is a square, *i.e.*, $b = a = \pi$. Then from (9) we have

$$\lambda_{m,n} = (m^2 + n^2). \tag{12}$$

Interchanging m and n leaves the eigen-value unchanged, but changes the eigen-function, thus

$$v_1 = \sin mx \sin ny$$

$$v_2 = \sin nx \sin my$$

generally differ from each other, but both of them belong to the same eigen-value $(m^2 + n^2)$.

Consequently, with the exception of the fundamental vibration $m = 1$, $n = 1$, and the overtones $m = n$ which are harmonic to it, each eigen-value of the square membrane is at least two-fold. This case is therefore degenerate.

9 · 3. *Perturbation Theory for the Non-Degenerate Case.*

The wave-equation for the unperturbed dynamical system has been found to be :

$$(\mathrm{H} - \mathrm{E})\psi = 0, \tag{1}$$

where H is the Hamiltonian of this undisturbed system standing for the operator $\mathrm{H}\left(q_r, \frac{h}{2\pi i}\frac{\partial}{\partial q_r}\right)$, and E is the eigen-value parameter.

Suppose we know the eigen-values and eigen-functions of (1) :

$$\mathrm{E}_n \ (n = 1, 2, 3, \cdots),$$

$$\psi_n \ (n = 1, 2, 3, \cdots).$$

Suppose further that each of the eigen-values ψ_n is simple, and that the eigen-functions are orthogonal and normalised :

$$\int \psi_m^* \psi_n \, dq = 0 \quad (m \neq n), \tag{2}$$

and

$$\int \psi_n^* \psi_n \, dq = \int \psi_n^2 \, dq = 1. \tag{3}$$

Let a small perturbation term $\epsilon A \psi$ be added now to the equation (1) on the right-hand side,

$$(H - E)\,\psi = \epsilon A \psi, \tag{4}$$

where ϵ is a small numerical factor, and A is a function of the q's, which is continuous in the whole domain. We wish to find out how the new eigen-values $\bar{E}_n$ and the new eigen-functions $\bar{\psi}_n$ of the equation (4) stand to the old eigen-values and eigen-functions of the equation (1).

From the continuity properties mentioned at the end of § 9·1, it is evident anyhow that for sufficiently small ϵ, $\bar{E}_n$ must lie in the neighbourhood of E_n, and $\bar{\psi}_n$ must lie in the neighbourhood of ψ_n, so that for a first approximation we may write for all $n \geqslant 1$:

$$\bar{E}_n = E_n + \epsilon\,\mu_n, \tag{5}$$

and

$$\bar{\psi}_n = \psi_n + \epsilon v_n. \tag{6}$$

Substituting (5) and (6) in (4), we get

$$(H - E_n - \epsilon\mu_n)\,(\psi_n + \epsilon v_n) = \epsilon A\,(\psi_n + \epsilon v_n),$$

or, neglecting the terms with ϵ^2:

$$(H - E_n)\,\psi_n + \epsilon\,\{(H - E_n)\,v_n - \mu_n\,\psi_n\} = \epsilon\,A\,\psi_n.$$

But, on account of (1), $(H - E_n)\,\psi_n = 0$, so that after division by ϵ, and transposition, the last equation reduces to

$$(H - E_n)\,v_n = (A + \mu_n)\,\psi_n. \tag{7}$$

The change v_n in the eigen-function is thus determined from the equation (7). The equation (7) is a non-homogeneous equation corresponding to the homogeneous equation (1), with the eigen-value E_n written for E. As remarked in (11) § 9·2, the equation (7) possesses a solution when, and only when, its right-hand side is orthogonal to the eigen-function ψ_n corresponding to the left-hand expression equated to zero. Thus we have

$$\int \psi_n^*\,(A + \mu_n)\,\psi_n\,dq = 0,$$

or,

$$\int \mu_n \psi_n^* \psi_n\,dq = -\int A\psi_n^2\,dq,$$

or, on account of (3):

$$\mu_n = -\frac{\int \psi_n^* A\,\psi_n\,dq}{\int \psi_n^* \psi_n\,dq} = -\int \psi_n^* A\,\psi_n\,dq. \tag{8}$$

The equation (8) gives the first order change μ_n in the eigen-value E_n due to the perturbation. Now, as the eigen-value E_n of (1) signifies mechanical energy, and the eigen-function ψ_n is comparable to motion with energy E_n, then the equation (8) gives the complete parallel to the well-known theorem in the perturbation theory of classical mechanics, that the perturbation of the energy, to a first approximation, is equal to the perturbing function averaged over the unperturbed motion.

We have to determine now the perturbation term v_n in the eigen-function. We can now suppose that the value of μ_n is known in the equation (7). We expand both v_n and $(A + \mu_n)\psi_n$ in a series of eigen-functions ψ_k of the unperturbed problem :

$$v_n = \sum_k v_{kn} \psi_k, \tag{9}$$

and

$$(A + \mu_n) \psi_n = \sum_k a_{kn} \psi_k, \tag{10}$$

where, on account of (15) § 8·6 (2),

$$a_{kn} = \int \psi_k^* (A + \mu_n) \psi_n \, dq.$$

Thus we see from (8) and (2) that

$$\left.\begin{aligned} a_{kn} &= \int \psi_k^* A \psi_n \, dq \quad \text{for } k \neq n, \\ &= 0 \qquad\qquad\qquad \text{for } k = n. \end{aligned}\right\} \tag{11}$$

Similarly, for v_{kn}, we would have used the integral

$$v_{kn} = \int \psi_k^* v_n \, dq, \tag{12}$$

but since we do not know v_n, this is useless for our purpose.

If we substitute (9) and (10) in (7), we get

$$\sum_k (H - E_n) v_{kn} \psi_k = \sum_k a_{kn} \psi_k.$$

But, for each $k \geqslant 1$ we have $(H - E_k) \psi_k = 0$, therefore

$$H \psi_k = E_k \psi_k.$$

Substituting this in the last equation we get

$$\sum_k v_{kn} (E_k - E_n) \psi_k = \sum_k a_{kn} \psi_k. \tag{13}$$

Since ψ_k is an independent set we can equate the coefficients of each ψ_k, and get for all $k \geqslant 1$:

$$v_{kn} (E_k - E_n) = a_{kn}, \tag{14}$$

or, if $k \neq n$, so that $E_k \neq E_n$, then

$$v_{kn} = \frac{a_{kn}}{E_k - E_n} \quad (k \neq n). \tag{15}$$

The coefficient v_{nn} remains indeterminate, because for $k = n$, $E_k = E_n$, and we cannot divide (14) by $E_k - E_n$. To determine v_{nn}, we use the still available condition that the new eigenfunctions $\bar{\psi}_n$ must be normalised, *i.e.*, that for all n

$$\int \bar{\psi}_n^* \bar{\psi}_n \, dq = 1. \tag{16}$$

Substituting (6) in this integral, we have:

$$\int (\psi_n + \epsilon v_n)^* (\psi_n + \epsilon v_n) \, dq = 1;$$

or, neglecting ϵ^2,

$$\int \psi_n^* \psi_n \, dq + \epsilon \int v_n^* \psi_n \, dq + \epsilon \int \psi_n^* v_n \, dq = 1.$$

But since the old ψ_n are normalised, therefore

$$\int v_n^* \psi_n \, dq + \int \psi_n^* v_n \, dq = 0. \tag{16 a}$$

Substituting (9) in (17) we get

$$\sum_k v_{kn}^* \int \psi_k^* \psi_n \, dq + \sum_k v_{kn} \int \psi_n^* \psi_k \, dq = 0. \tag{17}$$

Since for $k \neq n$ the integrals are already zero on account of (2), and the integrals do not vanish for $k = n$ on account of (3), we see that in this latter case, *viz.*, for $k = n$, the coefficients in (17) must vanish. Therefore for all n:

$$v_{nn} = v_{nn}^* = 0. \tag{18}$$

Thus, we have determined all the coefficients in the expansion (9), and substituting (9), (11), (15) in the equation (6), we have for all $n \geqslant 1$:

$$\begin{aligned} \psi_n &= \psi_n + \epsilon \sum_k{}' v_{kn} \psi_k \\ &= \psi_n + \epsilon \sum_k{}' \psi_k \frac{a_{kn}}{E_k - E_n} \\ &= \psi_n + \epsilon \sum_k{}' \psi_k \frac{\int A \psi_k^* \psi_n \, dq}{E_k - E_n}, \end{aligned} \tag{19}$$

where $\sum_k{}'$ means that the sum is to be extended from $k = 1$ to ∞ except $k = n$.

Similarly, from (5) and (8) we have for the new eigen-value

$$\bar{E}_n = E_n - \epsilon \int \psi_n^* A \, \psi_n \, dq. \tag{20}$$

$$(n = 1, 2, 3, \cdots)$$

It is necessary in every case to verify that the functions (19) and the values (20) really satisfy the wave-equation (4) for the perturbed system.

After determining this first approximation, we can determine the second approximation

$$\bar{E}_n = E_n + \epsilon\mu_n + \epsilon^2 \nu_n, \tag{21}$$

$$\bar{\psi}_n = \psi_n + \epsilon\, v_n + \epsilon^2 w_n, \tag{22}$$

in the following way.

Substituting (21) and (22) in (4), and taking account of the equations (1) and (7), and neglecting ϵ^3 and higher powers, we get on equating the coefficients of ϵ^2 on both sides:

$$(H - E_n)\, w_n = A\, v_n + \mu_n v_n + \nu_n \psi_n. \tag{23}$$

As remarked above, this non-homogeneous equation has a solution when, and only when, the right-hand side is orthogonal to the eigen-function of the left-hand side, *i.e.*, when

$$\int \psi_n^* \{(A + \mu_n)\, v_n + \nu_n \psi_n\}\, dq = 0, \tag{23a}$$

or

$$\begin{aligned} \nu_n \int \psi_n^* \psi_n\, dq &= -\int \psi_n^* (A + \mu_n)\, v_n\, dq \\ &= -\int (A + \mu_n)\, \psi_n^* \sum_k v_{kn} \psi_k\, dq \\ &= -\sum_k v_{kn} \int (A + \mu_n)\, \psi_n^* \psi_k\, dq \\ &= -\sum_k v_{kn}\, a_{nk}, \end{aligned}$$

on account of (11) and (18). Thus we get

$$\nu_n = -\sum_k a_{nk}\, v_{kn}. \tag{24}$$

Now suppose

$$w_n = \sum_k w_{kn}\, \psi_k, \tag{25}$$

and

$$(A + \mu_n)\, v_n + \nu_n\, \psi_n = \sum_k c_{kn}\, \psi_k, \tag{26}$$

where

$$\left.\begin{aligned} c_{kn} &= \int \psi_k^* \{(A + \mu_n)\, v_n + \nu_n \psi_n\}\, dq \\ &= 0 \qquad \text{for } k = n, \\ &= \int \psi_k^* (A + \mu_n)\, v_n\, dq, \text{ for } k \neq n, \end{aligned}\right\} \tag{27}$$

on account of (23 *a*) and (2). Therefore, if $k \neq n$, we have from (9):

$$
\begin{aligned}
c_{kn} &= \int \psi_k^* (A + \mu_n) \{\sum_l v_{ln}\, \psi_l\}\, dq \\
&= \sum_l v_{ln} \int A\psi_k^*\psi_l\, dq + \sum_l \mu_n v_{ln} \int \psi_k^*\psi_l\, dq \\
&= \sum_l v_{ln}\, a_{kl} + \mu_n\, v_{kn}, \qquad (28)
\end{aligned}
$$

on account of (11), (2) and (3)

Substituting (25) and (26) in (23), we get:

$$(H - E_n) \sum_k w_{kn}\psi_k = \sum_k c_{kn}\, \psi_k.$$

or

$$\sum_k w_{kn}\, (H - E_n)\, \psi_k = \sum_k c_{kn}\, \psi_k. \qquad (29)$$

But from (1), $H\, \psi_k = E_k\, \psi_k$, so that from (29) we get for all $k \geqslant 1$:

$$w_{kn}\, (E_k - E_n) = c_{kn}. \qquad (30)$$

If $k \neq n$, $E_k \neq E_n$ (non-degenerate case), therefore

$$
\begin{aligned}
w_{kn} &= \frac{c_{kn}}{E_k - E_n} \\
&= \frac{1}{E_k - E_n}\left\{\sum_{l=1}^{\infty} a_{kl}\, v_{ln} + \mu_n\, v_{kn}\right\}. \qquad (31)
\end{aligned}
$$

The formula (30) does not give us the diagonal coefficients w_{nn}. To determine these, we use again the condition that the new eigen-functions must be normalised. Thus substituting (22) in (16):

$$\int(\psi_n^* + \epsilon v_n^* + \epsilon^2 w_n^*)\, (\psi_n + \epsilon v_n + \epsilon^2 w_n)\, dq = 1,$$

or

$$
\begin{aligned}
&\int\psi_n^*\psi_n\, dq + \epsilon \int v_n^*\psi_n + \psi_n^* v_n\, dq \\
&+ \epsilon^2 \int \{(w_n^*\psi_n + \psi_n^* w_n) + v_n^* v_n\}\, dq = 1, \qquad (32)
\end{aligned}
$$

neglecting ϵ^3 and higher powers. But on account of (3) $\int \psi_n^*\psi_n\, dq = 1$, and on account of (16 *a*), $\int(v_n^*\, \psi_n + \psi_n^*\, v_n)\, dq = 0$, so that from (32) we get

$$\int (w_n^*\psi_n + \psi_n^* w_n)\, dq = -\int v_n^* v_n\, dq,$$

or

$$\int\{\psi_n \sum_k w_{kn}^*\, \psi_k^* + \psi_n^* \sum_k w_{kn}\psi_k\}\, dq = -\int \{\sum_k v_{kn}\psi_k\}\, \{\sum_l v_{ln}^*\psi_l\}\, dq,$$

or

$$2\,w_{nn} = -\sum_k v_{kn}^2,$$

or

$$w_{nn} = -\tfrac{1}{2}\sum_k v_{kn}^2. \quad (n = 1, 2, \cdots) \quad (33)$$

Thus we have determined completely the second approximation. Subsequently, we can determine the third, fourth and higher approximations.

9·4. *The Atom under an Electric Field—The Stark-effect.*

As an example of the perturbation method we consider the Stark-effect produced by the application of an external electric field to the atom. In the year 1913, J. Stark demonstrated experimentally that the spectrum of hydrogen in an intense electric field is much more complicated than under ordinary conditions. He found, in fact, that each line of the Balmer series is split up into a certain number of components, the number increasing with the increasing frequency of the line. The components separate out symmetrically on either side of the original line. If the distance of the least displaced component from the centre of the system is Δ, then the distances of the other components from the centre are integral multiples of Δ, and Δ itself is proportional to the strength of the field.

The classical electrodynamics is quite unable to account for this Stark-effect. A complete explanation on Bohr's quantum theory, accounting for all the detailed features of the effect, was given by K. Schwarzschild and P. Epstein in 1916. On wave mechanics it was solved by Schrödinger himself in 1926, in the third memoire in which he developed the perturbation theory.

Suppose an electric field of intensity F is acting on the hydrogen atom in the direction of the z-axis, then the potential energy V of the system is given by

$$V = -\frac{e^2}{r} + Fez,$$

where the origin is taken at the nucleus which is considered to be at rest. The relativistic (fine-structure) separation is so small compared to the effect due to the field, that it may be neglected for a first approximation.

Substituting the above value of V in the wave-equation, we get

$$\nabla^2 \psi + \frac{8\pi^2 m_0}{h^2}\left(\mathrm{E} + \frac{e^2}{r} - \mathrm{F}e\, z\right)\psi = 0, \tag{1}$$

so that the perturbation term is $-\mathrm{F}ez\psi$. This perturbation term is added to the hydrogen atom treated in § 8·5. Following the method given in the last section, we must expand this perturbation term in a series of eigen-functions of (2) § 8·5, and proceed to find the perturbed energy values. But the problem becomes easier if we transform first to the parabolic co-ordinates, and then apply the perturbation method. This latter procedure was adopted by Epstein in the older quantum theory.

Thus we write

$$\begin{aligned} x &= \sqrt{\xi\eta}\ \cos\phi, \\ y &= \sqrt{\xi\eta}\ \sin\phi, \\ z &= \tfrac{1}{2}(\xi - \eta), \end{aligned} \tag{2}$$

the domain of the new co-ordinates being

$$0 \leqslant \xi < \infty, \quad 0 \leqslant \eta < \infty, \quad 0 \leqslant \phi \leqslant 2\pi. \tag{3}$$

The Jacobian of the transformation can be easily calculated to be

$$\frac{\partial(x, y, z)}{\partial(\xi, \eta, \phi)} = \begin{vmatrix} \dfrac{\partial x}{\partial \xi} & \dfrac{\partial y}{\partial \xi} & \dfrac{\partial z}{\partial \xi} \\ \dfrac{\partial x}{\partial \eta} & \dfrac{\partial y}{\partial \eta} & \dfrac{\partial z}{\partial \eta} \\ \dfrac{\partial x}{\partial \phi} & \dfrac{\partial y}{\partial \phi} & \dfrac{\partial z}{\partial \phi} \end{vmatrix} = \tfrac{1}{4}(\xi + \eta), \tag{4}$$

and the space element is therefore

$$dx\, dy\, dz = \frac{\partial(x, y, z)}{\partial(\xi, \eta, \phi)}\, d\xi\, d\eta\, d\phi = \tfrac{1}{4}(\xi + \eta)\, d\xi\, d\eta\, d\phi. \tag{5}$$

Further, we have

$$r^2 = x^2 + y^2 + z^2 = \xi\eta + \tfrac{1}{4}(\xi - \eta)^2 = \tfrac{1}{4}(\xi + \eta)^2. \tag{6}$$

Substituting these values in (1) we easily find that the wave-equation in the new variables becomes :

$$\frac{\partial}{\partial \xi}\left(\xi \frac{\partial \psi}{\partial \xi}\right) + \frac{\partial}{\partial \eta}\left(\eta \frac{\partial \psi}{\partial \eta}\right) + \tfrac{1}{4}\left(\frac{1}{\xi} + \frac{1}{\eta}\right)\frac{\partial^2 \psi}{\partial \phi^2}$$

$$+ \frac{2\pi^2 m_0}{h^2}\{\mathrm{E}(\xi + \eta) + 2e^2 - \tfrac{1}{2}e\mathrm{F}(\xi^2 - \eta^2)\}\psi = 0. \tag{7}$$

The co-ordinate ϕ does not occur explicitly in (7), and therefore ϕ will enter into the solution as cos $m\phi$ or sin $m\phi$. Since the solution is to be continuous and single-valued, we must have

$$m = 0, 1, 2, 3, \cdots \tag{8}$$

We therefore assume that the solution of (7) is

$$\psi = f(\xi)\, g(\eta) \frac{\cos}{\sin} m\phi, \tag{9}$$

and substituting in (7) get the two equations

$$\frac{\partial}{\partial\xi}\left(\xi\frac{df}{d\xi}\right) + \frac{2\pi^2 m_0}{h^2}\left(-\tfrac{1}{2}eF\xi^2 + E\xi + e^2 - \lambda - \frac{m^2h^2}{8\pi^2 m_0}\frac{1}{\xi}\right)f = 0, \tag{10}$$

$$\frac{\partial}{\partial\eta}\left(\eta\frac{dg}{d\eta}\right) + \frac{2\pi^2 m_0}{h^2}\left(\tfrac{1}{2}eF\eta^2 + E\eta + e^2 + \lambda - \frac{m^2h^2}{8\pi^2 m_0}\frac{1}{\eta}\right)g = 0, \tag{11}$$

where λ is another constant which is the eigen-value parameter.

Both of the equations (10) and (11) are of the type

$$\frac{\partial}{\partial\zeta}\left(\zeta\frac{dX}{d\zeta}\right) + \left(D\zeta^2 + A\zeta + 2B + \frac{C}{\zeta}\right)X = 0, \tag{12}$$

where, for (10) we have

$$A = \frac{2\pi^2 m_0 E}{h^2},\ B_1 = \frac{\pi^2 m_0}{h^2}(e^2 - \lambda),\ C = -\frac{m^2}{4},\ D_1 = -\frac{\pi^2 m_0 eF}{h^2}, \tag{13}$$

and, for (11) we have

$$A = \frac{2\pi^2 m_0 E}{h^2},\ B_2 = \frac{\pi^2 m_0}{h^2}(e^2 + \lambda),\ C = -\frac{m^2}{4},\ D_2 = \frac{\pi^2 m_0 eF}{h^2}. \tag{14}$$

For $D = 0$, *i.e.*, $F = 0$, the equation (12) reduces to:

$$\frac{\partial}{\partial\zeta}\left(\zeta\frac{\partial X}{\partial\zeta}\right) + \left(A\zeta + 2B + \frac{C}{\zeta}\right)X = 0, \tag{15}$$

which is the equation of the simple hydrogen atom expressed in parabolic co-ordinates.

Now we transform the equations (12) and (15) by

$$X = \zeta^{\frac{m}{2}}\, u, \tag{16}$$

and then change the independent variable from ζ to θ by

$$2\zeta\sqrt{-A} = \theta. \tag{17}$$

Then (12) is transformed to

$$\frac{d^2u}{d\theta^2} + \frac{m+1}{\theta}\frac{du}{d\theta} + \left\{\frac{D}{(2\sqrt{-A})^3}\theta - \tfrac{1}{4} + \frac{B}{\sqrt{-A}}\frac{1}{\theta}\right\}u = 0, \tag{18}$$

and (15) is transformed to

$$\frac{d^2u}{d\theta^2} + \frac{m+1}{\theta}\frac{du}{d\theta} + \left\{-\tfrac{1}{4} + \frac{B}{\sqrt{-A}}\frac{1}{\theta}\right\}u = 0. \tag{19}$$

In the theory of differential equations, the polynomial which satisfies the differential equation

$$x\frac{d^2y}{dx^2} + (1-x)\frac{dy}{dx} + ky = 0, \tag{20}$$

where k is an integer, is called the kth Laguerre Polynomial, and is denoted by $L_k(x)$. If we first replace k by $m+k$, and then differentiate m times, we find that the mth derivative of the $(m+k)$th Laguerre Polynomial, which is denoted by $L^m_{m+k}(x)$, satisfies

$$x\frac{d^2y}{dx^2} + (m+1-x)\frac{dy}{dx} + ky = 0. \tag{21}$$

If we transform (21) by writing $y = e^{-\frac{x}{2}} L^m_{m+k}(x)$, we get the equation

$$\frac{d^2y}{dx^2} + \frac{m+1}{x}\frac{dy}{dx} + \left\{-\tfrac{1}{4} + \left(k + \frac{m+1}{2}\right)\frac{1}{x}\right\}y = 0. \tag{22}$$

Suppose that m is a fixed integer in the equation (22), while k is the eigen-value parameter. Then (22) has the eigen-functions $e^{-\frac{x}{2}} L^m_{m+k}(x)$ corresponding to the eigen-values $k = 0, 1, 2, 3, \cdots$ Comparing (19) and (22) we find that the eigen-values of (19) are

$$\frac{B}{\sqrt{-A}} = k + \frac{m+1}{2}, \quad (k = 0, 1, 2, 3, \cdots), \tag{23}$$

and the corresponding eigen-functions are

$$u_k(\theta) = e^{-\frac{\theta}{2}} L^m_{m+k}(\theta). \tag{24}$$

We have now to calculate the first approximation to the perturbed eigen-values. The perturbation is caused by including the D-term in (18). To make the equation (18) self-adjoint, we must multiply it by θ^{m+1} throughout, so that it becomes:

$$\theta^{m+1}\frac{d^2u}{d\theta^2} + (m+1)\,\theta^m\frac{du}{d\theta} +$$

$$\left\{\frac{D}{(2\sqrt{-A})^3}\,\theta^{m+2} - \tfrac{1}{4}\,\theta^{m+1} + \frac{B}{\sqrt{-A}}\,\theta^m\right\}u = 0,$$

or

$$\frac{d}{d\theta}\left\{\theta^{m+1}\frac{du}{d\theta}\right\}+\left\{\frac{D}{(2\sqrt{-A})^3}\theta^{m+2}-\tfrac{1}{4}\theta^{m+1}+\frac{B}{\sqrt{-A}}\theta^{m}\right\}u=0, \tag{25}$$

which is in the self-adjoint form.

The perturbing function, which was denoted by ϵA in (4) § 9· 3, is $\frac{D\theta^{m+2}}{(2\sqrt{-A})^3}$ in (25). Moreover, in that equation the eigen-value parameter E had a factor unity only, whereas in (25) the eigen-value parameter has a factor θ^m, which is called the density function, and is usually denoted by ρ (θ). The formulæ (5) and (8) of § 9·3, give therefore for the perturbation $\epsilon_k = \epsilon\,\mu_k$ of the *k*th eigen-value*:

$$\epsilon_k = -\frac{\int \epsilon\, A u_k^2\, dq}{\int \rho\, u_k^2\, dq} = -\frac{D}{(2\sqrt{-A})^3}\;\frac{\int_0^\infty \theta^{m+2}e^{-\theta}\{L_{m+k}^{m}(\theta)\}^2\,d\theta}{\int_0^\infty \theta^{m}e^{-\theta}\{L_{m+k}^{m}(\theta)\}^2\,d\theta}. \tag{26}$$

It is known from the theory of the Laguerre Polynomials, that

$$\int_0^\infty \theta^{m+2}e^{-\theta}\{L_{m+k}^{m}(\theta)\}^2\,d\theta = \frac{\{(m+k)\,!\}^3}{k!},$$

$$\int_0^\infty \theta^{m}e^{-\theta}\{L_{m+k}^{m}(\theta)\}^2\,d\theta$$

$$= \frac{\{(m+k)\,!\}^3}{k\,!}\Big/(m^2+6mk+6k^2+6k+3m+2).$$

Substituting these values in (26), we get

$$\epsilon_k = -\frac{D}{(2\sqrt{-A})^3}(m^2+6mk+6k^2+6k+3m+2). \tag{27}$$

The *k*th perturbed eigen-value of (25) is therefore, from (23) and (27) :

$$\frac{B}{\sqrt{-A}} = \frac{m+1}{2}+k+\epsilon_k. \tag{28}$$

Applying this result to the equation (10), with the constants given by (13), we get

$$\frac{B_1}{\sqrt{-A}} = \frac{m+1}{2}+k_1+\epsilon_{k_1}. \tag{29}$$

---

* Since the eigen-functions are real, the conjugate complex is equal to the function itself.

Similarly, applying (28) to the equation (11), with the constants given by (14), we obtain

$$\frac{B_2}{\sqrt{-A}} = \frac{m+1}{2} + k_2 + \epsilon_{k_2}. \tag{30}$$

Adding (29) and (30) :

$$\begin{aligned}\frac{1}{\sqrt{-A}}(B_1 + B_2) &= m + 1 + k_1 + k_2 + \epsilon_{k_1} + \epsilon_{k_2} \\ &= n + \epsilon_{k_1} + \epsilon_{k_2},\end{aligned} \tag{31}$$

where the principle quantum number n is given by :

$$n = m + k_1 + k_2 + 1. \tag{32}$$

Thus from (31) we get on squaring and solving for A :

$$\begin{aligned}A &= -\frac{(B_1 + B_2)^2}{(n + \epsilon_{k_1} + \epsilon_{k_2})^2} \\ &= -\frac{(B_1 + B_2)^2}{n^2}\left\{1 - \frac{2}{n}(\epsilon_{k_1} + \epsilon_{k_2})\right\},\end{aligned} \tag{33}$$

to the first approximation, since ϵ_{k_1} and ϵ_{k_2} are small. Further, to the approximation we are aiming at, we can use the unperturbed value of A in calculating ϵ_{k_1} and ϵ_{k_2}. Since for the unperturbed hydrogen atom $E = -\frac{2\pi^2 m_0 e^4}{h^2 n^2}$, therefore on account of (13), the unperturbed value of A is :

$$A = \frac{2\pi^2 m_0 E}{h^2} = \frac{2\pi^2 m_0}{h^2}\left(-\frac{2\pi^2 m_0 e^4}{h^2 n^2}\right) = -\frac{4\pi^4 m_0^2 e^4}{h^4 n^2},$$

so that

$$(2\sqrt{-A})^3 = \left\{2 \cdot \frac{2\pi^2 m_0 e^2}{h^2 n}\right\}^3 = \frac{64\pi^6 m_0^3 e^6}{h^6 n^3}. \tag{34}$$

Substituting this value and the values of D_1, D_2 from (13) and (14) in the equation (27), we obtain

$$\epsilon_{k_1} = \frac{Fh^4 n^3}{64\pi^4 m_0^2 e^5}(m^2 + 6mk_1 + 6k_1^2 + 6k_1 + 3m + 2),$$

$$\epsilon_{k_2} = -\frac{Fh^4 n^3}{64\pi^4 m_0^2 e^5}(m^2 + 6mk_2 + 6k_2^2 + 6k_2 + 3m + 2),$$

so that

$$\begin{aligned}\epsilon_{k_1} + \epsilon_{k_2} &= \frac{6Fh^4 n^3}{64\pi^4 m_0^2 e^5}\left\{(k_1 - k_2)(m + k_1 + k_2 + 1)\right\} \\ &= \frac{3Fh^4 n^4}{32\pi^4 m_0^2 e^5}(k_1 - k_2),\end{aligned} \tag{35}$$

on account of (32). From (13) and (14) we find also

$$B_1 + B_2 = \frac{\pi^2 m_0}{h^2}\left\{(e^2 - \lambda) + (e^2 + \lambda)\right\} = \frac{2\pi^2 m_0 e^2}{h^2}. \tag{36}$$

Substituting (35) and (36) in (33), we get finally

$$A = \frac{2\pi^2 m_0 E}{h^2} = -\left(\frac{2\pi^2 m_0 e^2}{h^2}\right)^2 \frac{1}{n^2}\left\{1 - \frac{2}{n}\,\frac{3Fh^4 n^4}{32\pi^4 m_0^2 e^5}(k_1 - k_2)\right\},$$

$$E = -\frac{2\pi^2 m_0 e^4}{h^2 n^2} - \frac{3}{8}\,\frac{Fh^2 n\,(k_2 - k_1)}{\pi^2 m_0 e}. \tag{37}$$

This is the first order approximation for the Stark-effect, and the formula (37) is the same as that derived by Epstein on Bohr's quantum theory. Epstein's quantum number k is now replaced by the number $k_2 - k_1$. Moreover, the numbers k_1, k_2 and m can take the value zero, but the principal quantum number n, being equal to $m + k_1 + k_2 + 1$, cannot take the value zero. The value zero is thus automatically excluded in the new quantum mechanics. Special artifices were required for this purpose in Bohir's theory.

Writing $W = -E$, and

$$k = k_2 - k_1, \tag{38}$$

we get from (37)

$$W = \frac{2\pi^2 m_0 e^4}{h^2 n^2} + \frac{3Fh^2}{8\pi^2 m_0 e}\,nk. \tag{39}$$

Equation (32) shows that $k_1 + k_2$ $(= n - 1 - m)$, must always be less than n, so that from (38) :

$$|k| < n \tag{40},$$

i.e., for any fixed n, k can take only the following values :

$$k = -(n-1), \cdots, -1, 0, +1, \cdots +(n-1). \tag{41}$$

The frequency ν of the transition $n \rightarrow n'$, $k \rightarrow k'$ is given by

$$\nu = \frac{W' - W}{h} = \frac{2\pi^2 m_0 e^4}{h^3}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right) + \frac{3Fh}{8\pi^2 m_0 e}(n'k' - nk). \tag{42}$$

The frequency ν_0 of the original line in the absence of the field is

$$\nu_0 = \frac{2\pi^2 m_0 e^4}{h^3}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right),$$

so that if we put

$$\Delta = \frac{3Fh}{8\pi^2 m_0 e}, \tag{43}$$

we get

$$\nu - \nu_0 = \Delta\,(n'k' - nk). \tag{44}$$

From (41) we know that for a given n and n', k and k' can have the following $(2n - 1)$ and $(2\,n' - 1)$ values respectively :

$$k = -(n-1),\ -(n-2),\ \cdots,\ -1,\ 0,\ +1,\ \cdots,\ (n-2),\ (n-1),$$

$$k' = -(n'-1),\ -(n'-2),\ \cdots,\ -1,\ 0,\ +1,\ \cdots,\ (n'-2),\ (n'-1)\ ;$$

and since each of the k can combine with each of the k', we get altogether $(2n - 1)(2n' - 1)$ components for the line $n \longrightarrow n'$. Thus for the H_α line of the Balmer series, $n = 3$, $n' = 2$, so that (44) becomes :

$$\nu - \nu_0 = \Delta\,(2k' - 3k). \tag{45}$$

Now since $|k| < 3$, it can have only the five values $-2, -1, 0, 1, 2$; and since $|k'| < 2$, it can have only the three values $-1, 0, +1$. Taking each k with each k', we get $5 \times 3 = 15$ values of $2k' - 3k$:

$$2k' - 3k = -8,\ -6,\ -5,\ -4,\ -3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 8.$$

Thus the H_α line is resolved into (fifteen) components, arranged on both sides of the original line, and separated by the distance Δ. These results are in good agreement with those given by experiment.

9·5. *Perturbation Theory for the Degenerate Case.*

We remarked in § 9·2 that it often happens that one eigen-value has several eigen-functions, in which case the problem is said to be degenerate.

An eigen-value E_k is called α-fold when it possesses α linearly independent eigen-functions :

$$\psi_{k1},\ \psi_{k2},\ \cdots,\ \psi_{ki},\ \cdots\ \psi_{k\alpha}. \tag{1}$$

Linear independence means that no numbers $c_1, c_2, \cdots, c_\alpha$ exist, which are not all zero, and for which the relation

$$\sum_{i=1}^{\alpha} c_i\ \psi_{ki} = 0 \tag{2}$$

holds. The wave-equation

$$(H - E_k)\ \psi = 0 \tag{3}$$

has therefore the α-solutions (1). Then it is easy to see that each of these α-functions (1) is orthogonal to each of the other eigen-functions belonging to another eigen-value $E_{k'}$ ($k' \neq k$). On the contrary, these α-functions are not, in general, orthogonal to one another. But, by suitable linear transformations we can always bring about this mutual orthogonality. We shall, therefore, assume that these ψ_{ki} have been normalised to 1, and have been made orthogonal not only towards the other eigen-functions $\psi_{k'}$, but also among themselves.

But every linear combination of the solutions (1) of a linear equation (3) is itself a solution, and since an orthogonal substitution leaves the mutual orthogonality unaltered, we see that the functions (1) are still not uniquely defined. Rather, we can submit them to any arbitrary orthogonal transformation without disturbing the character of their normalisation and orthogonality. In fact, if we set:

$$\psi_{kl} = \sum_{i=1}^{\alpha} \beta_{li}\, \psi_{ki}, \tag{4}$$

where β_{li} are the coefficients of any arbitrary orthogonal transformation, characterised by

$$\sum_{i} \beta_{l'i}^{*}\, \beta_{li} = \delta_{ll'} = \begin{cases} 0 & \text{if } l \neq l' \\ 1 & \text{if } l = l', \end{cases} \tag{5}$$

and if

$$\psi_{kl'} = \sum_{j=1}^{\alpha} \beta_{l'j}\, \psi_{kj},$$

then

$$\int \psi_{kl'}^{*}\, \psi_{kl}\, dq = \sum_{i=1}^{\alpha} \sum_{j=1}^{\alpha} \beta_{l'j}^{*}\, \beta_{li} \int \psi_{kj}^{*}\, \psi_{ki}\, dq. \tag{6}$$

But, since the functions (1) are normalised and orthogonal:

$$\int \psi_{kj}^{*}\, \psi_{ki}\, dq\, \delta_{lj} = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j, \end{cases} \tag{7}$$

therefore substituting (7) in (6) we get on account of (5):

$$\int \psi_{kl'}^{*}\, \psi_{kl}\, dq = \sum_{i=1}^{\alpha} \beta_{l'i}^{*}\, \beta_{li} \tag{8}$$
$$= \delta_{ll'},$$

showing that the new functions ψ_{kl} are also orthogonal and normalised.

The coefficients β_{li} of the orthogonal transformation (4) are arbitrary. It is of this very arbitrariness that we make use in the perturbation theory.

As in the previous section, suppose the perturbed equation is

$$(H - E)\,\psi = \epsilon A\,\psi. \tag{9}$$

The perturbation would in general split the α-fold eigen-value E_k in α simple values $\bar{E}_{kl}$, to each of which would correspond a single eigen-function $\bar{\psi}_{kl}$, given by

$$\bar{E}_{kl} = E_k + \epsilon\,\mu_l,\quad (l = 1, 2, \cdots, \alpha), \tag{10}$$

and

$$\bar{\psi}_{kl} = \psi_{kl} + \epsilon v_l = \sum_{i=1}^{\alpha} \beta_{li}\,\psi_{ki} + \epsilon v_l. \tag{11}$$

The index k is of course fixed throughout as we are considering the perturbation of the eigen-value E_k only.

We select a definite perturbed eigen-value and eigen-function by giving a definite value to the index l in (10) and (11), and substitute them in (9):

$$(H - E_k - \epsilon\,\mu_l)\,\Big(\sum_{i=1}^{\alpha} \beta_{li}\,\psi_{ki} + \epsilon v_l\Big) = \epsilon A\,\Big(\sum_{i=1}^{\alpha} \beta_{li}\,\psi_{ki} + \epsilon v_l\Big),$$

or, neglecting ϵ^2 and transposing,

$$(H - E_k)\sum_{i=1}^{\alpha} \beta_{li}\psi_{ki} + \epsilon\,(H - E_k)\,v_l = \epsilon\,\Big\{A\sum_{i=1}^{\alpha} \beta_{li}\psi_{ki} + \mu_l\sum_{i=1}^{\alpha} \beta_{li}\,\psi_{ki}\Big\}.$$

For each of the ψ_{ki}, we have $(H - E_k)\,\psi_{ki} = 0$, therefore from the last equation we get on dividing up by ϵ, for all $l = 1, 2, \cdots \alpha$:

$$(H - E_k)\,v_l = \sum_{i=1}^{\alpha} \beta_{li}\,(A + \mu_l)\,\psi_{ki}. \tag{12}$$

This is again a non-homogeneous equation, whose solution, as remarked in the preceding sections, is only then possible when the right-hand side is orthogonal to all the eigen-functions of the homogeneous equation (3) belonging to E_k. We must have therefore:

$$\sum_{i=1}^{\alpha} \beta_{li} \int \psi_{km}^{*}\,(A + \mu_l)\,\psi_{ki}\,dq = 0, \tag{13}$$

$$(m = 1, 2, \cdots, \alpha).$$

On account of (8), this reduces to

$$\beta_{lm}\,\mu_l = -\sum_{i=1}^{\alpha} \beta_{li} \int \psi_{km}^{*}\,A\,\psi_{ki}\,dq. \tag{14}$$

$$(m = 1, 2, 3, \cdots, \alpha).$$

Let

$$a_{im} = - \int \psi_{km}^* \, \mathrm{A} \, \psi_{ki} \, dq, \tag{15}$$

$$(i, m = 1, 2, \cdots, \alpha)$$

then (14) reduces to a system of α linear homogeneous equations

$$\beta_{lm} \, \mu_l = \sum_{i=1}^{\alpha} \beta_{li} \, a_{im}, \tag{16}$$

$$(m = 1, 2, 3, \cdots, \alpha),$$

or, writing in the expanded form,

$$\begin{array}{lllll} \beta_{l1}(a_{11} - \mu_l) & + \beta_{l2} \, a_{21} + \beta_{l3} \, a_{31} & + \cdots & + \beta_{l\alpha} \, a_{\alpha 1} & = 0 \\ \beta_{l1} \, a_{12} & + \beta_{l2}(a_{22} - \mu_l) & + \cdots\cdots & + \beta_{l\alpha} \, a_{\alpha 2} & = 0 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \\ \beta_{l1} \, a_{1\alpha} & + \beta_{l2} \, a_{2\alpha} & + \cdots\cdots & + \beta_{l\alpha}(a_{\alpha\alpha} - \mu_l) & = 0. \end{array}$$

It is known from the theory of algebraic equations that these α homogeneous linear equations in the α unknown quantities $\beta_{l1}, \cdots, \beta_{l\alpha}$, have a non-trivial solution only when the determinant of the coefficients vanishes :

$$\begin{vmatrix} a_{11} - \mu_l & a_{12} & a_{13} \cdots a_{1\alpha} \\ a_{21} & a_{22} - \mu_l & a_{23} \cdots a_{2\alpha} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{\alpha 1} \cdots\cdots & \cdots\cdots & a_{\alpha\alpha} - \mu_l \end{vmatrix} = 0. \tag{17}$$

This is an equation of the degree α in μ_l, and it is shown in Algebra that it has α real roots. To each of these roots there belongs a special value system of β_{li}, and consequently a definite eigen-function ψ_{kl} of the family.

We thus get all the α-perturbations $\mu_l \, (l = 1, 2, \cdots, \alpha)$ of the eigen-value E_k simultaneously from the equation (17) provided this equation has no multiple roots, which is usually the case. These then give us the corresponding α eigen-functions of the unperturbed system.

To determine the eigen-functions $\bar{\psi}_{kl}$ of the perturbed equation (9) we have to solve the equation (12). For this purpose we expand v_l as a series of the complete set of eigen-functions of (3) :

$$v_l = \sum_{k'=0}^{\infty} \sum_{i'=1}^{\alpha} \gamma_{l,k'i'} \, \psi_{k'i'}. \tag{18}$$

Similarly, we develop the right-hand side expression in (12) in a complete set of ψ_{kl}:

$$\sum_{i=1}^{\alpha} \beta_{li}\,(A+\mu_l)\,\psi_{ki} = \sum_{k'=0}^{\infty}\sum_{i'=1}^{\alpha} C_{l,k'i'}\,\psi_{k'i'}, \tag{19}$$

where

$$C_{l,k'i'} = \sum_{i=1}^{\alpha} \beta_{li} \int \psi_{k'i'}^{*}\,(A+\mu_l)\,\psi_{ki}\,dq. \tag{20}$$

On account of the orthogonality of ψ_k and $\psi_{k'}$, and from the equation (13), we get

$$\begin{aligned} C_{l,k'i'} &= \sum_{i=1}^{\alpha} \beta_{li} \int \psi_{k'i'}\,A\,\psi_{ki}\,dq, \text{ for } k' \neq k. \\ &= 0, \text{ for } k' = k. \end{aligned} \tag{21}$$

Substituting (18) and (19) in (12), we get

$$\sum_{k'=0}^{\infty}\sum_{i'=1}^{\alpha} \gamma_{l,k'i'}\,(H-E_k)\,\psi_{k'i'} = \sum_{k'=0}^{\infty}\sum_{i'=0}^{\alpha} C_{l,k'i'}\,\psi_{k'i'}. \tag{22}$$

But H $\psi_{k'i'} = E_{k'}\,\psi_{k'i'}$, therefore

$$\sum_{k'}\sum_{i'} \gamma_{l,k'i'}\,(E_{k'}-E_k)\,\psi_{k'i'}, = \sum_{k'}\sum_{i'} C_{l,k'i'}\,\psi_{k'i'}. \tag{23}$$

If $k' \neq k$, then $E_{k'}$, $\neq$ E_k, so that equating the coefficients of $\psi_{k'i'}$ we get, on division,

$$\gamma_{l,k'i'} = \frac{C_{l,k'i'}}{E_{k'}-E_k} = \frac{1}{E_{k'}-E_k} \sum_{i=1}^{\alpha} \beta_{li} \int \psi_{k'i'}^{*}\,A\,\psi_{ki}\,dq. \tag{24}$$

For $k' = k$ the coefficients $\gamma_{l,k',i'}$ cannot be determined from (23). To determine them we employ the still available condition that the new eigen-functions $\psi_{k,l}$ should be normalised, *i.e.*,

$$\int \psi_{k,l}^{*}\,\psi_{k,l}\,dq = 1, \tag{25}$$

or, from (11),

$$\int (\psi_{kl}^{*} + \epsilon v_l^{*})\,(\psi_{k,l} + \epsilon v_l)\,dq = 1,$$

or, neglecting ϵ^2 and taking account of (8),

$$\int \psi_{k,l}^{*}\,v_l\,dq = 0.$$

Substituting in this from (18) we have

$$\sum_{k'}\sum_{k} \int \psi_{kl}^{*}\,\gamma_{l,k'i'}\,\psi_{k'i'}\,dq = 0. \tag{26}$$

For $k' \neq k$, the integral in (26) is zero on account of (8). For $k' = k$, therefore, $\gamma_{l,k'i'}$ should be zero. Thus for all i':

$$\gamma_{l,k'i'} = 0, \quad (k' = k). \tag{27}$$

We have thus determined all the coefficients $\gamma_{l,k'\iota'}$ and with them the function v_l. We have therefore, to the first approximation,

$$\begin{aligned} \psi_{kl} &= \psi_{kl} + \epsilon v_l \\ &= \sum_{\iota=1}^{a} \beta_{l\iota} \left\{ \psi_{k\iota} + \epsilon \sum_{k'=0}^{\infty}{}' \sum_{\iota'=1}^{a} \frac{\psi_{k'\iota'}}{E_{k'} - E_k} \int \psi_{k'\iota'}^{*} \, A \, \psi_{k\iota} \, dq \right\} \\ &\qquad (l = 1, 2, \cdots a), \end{aligned} \tag{28}$$

where $\sum_{k'}'$ indicates that all the terms with $k' = k$ are to be omitted.

The theory of this section has a very important application in the wave-mechanics of a system of similar particles, such as the neutral helium atom, the singly ionised lithium atom, etc.

9·6. *The Theory of the Neutral Helium Atom.*

Several attempts were made before 1925 to treat the neutral helium atom on the older quantum theory. But the results arrived at were hopelessly in contradiction with the facts of experiment. It was only after the advent of the new quantum mechanics that the problem could be successfully tackled, and it was Heisenberg who, in June 1926, gave a preliminary sketch of the problem of many bodies in quantum mechanics.

The neutral helium atom consists of two electrons revolving round a nucleus. Now, these two electrons are absolutely indistinguishable from each other, and no observable change is made when they are interchanged. These two observationally indistinguishable states are therefore counted as the same state in the new mechanics. This gives rise to the "resonance" phenomenon, which has no analogue in the classical theory.

Take the nucleus (of charge $+\mathrm{N}e$, $\mathrm{N} = 2$) as the origin, and suppose that one of the electrons is at the point (x_1, y_1, z_1), and the other electron is at the point (x_2, y_2, z_2). Let the distance of the first electron from the nucleus be r_1, that of the second electron from the nucleus be r_2, and the mutual distance of the two electrons be r_{12}. If the mass of the electron is denoted by μ, their velocities by (u_1, v_1, w_1), (u_2, v_2, w_2), and the kinetic, potential and total energies of the system by T, V and E respectively, then

$$\left. \begin{aligned} \mathrm{T} &= \tfrac{1}{2}\mu \left(u_1^2 + v_1^2 + w_1^2 + u_2^2 + v_2^2 + w_2^2 \right), \\ \mathrm{V} &= -\frac{\mathrm{N}e^2}{r_1} - \frac{\mathrm{N}e^2}{r_2} + \frac{e^2}{r_{12}}, \\ \mathrm{E} &= \mathrm{T} + \mathrm{V}. \end{aligned} \right\} \tag{1}$$

In order to make a transition from the corpuscular mechanics to wave-mechanics, we must take each momentum to be a differential operator:

$$\left.\begin{aligned} \mu u_1 &= \frac{h}{2\pi i}\frac{\partial}{\partial x_1}, \quad \mu v_1 = \frac{h}{2\pi i}\frac{\partial}{\partial y_1}, \quad \mu w_1 = \frac{h}{2\pi i}\frac{\partial}{\partial z_1}, \\ \mu u_2 &= \frac{h}{2\pi i}\frac{\partial}{\partial x_2}, \quad \mu v_2 = \frac{h}{2\pi i}\frac{\partial}{\partial y_2}, \quad \mu w_2 = \frac{h}{2\pi i}\frac{\partial}{\partial z_2} \end{aligned}\right\}. \tag{2}$$

Substituting these in $T = E - V$, and operating on a wave-function ψ, we get:

$$-\frac{h^2}{8\pi^2\mu}\left\{\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2}\right\}\psi = (E - V)\psi$$

or

$$\nabla^2\psi + \frac{8\pi^2\mu}{h^2}(E - V)\psi = 0, \tag{3}$$

where

$$\left.\begin{aligned} \nabla^2\psi &= \nabla_1^2\psi + \nabla_2^2\psi \\ \nabla_1^2\psi &= \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}\right)\psi \\ \nabla_2^2\psi &= \left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2}\right)\psi \end{aligned}\right\}. \tag{4}$$

To a first approximation, we neglect the interaction term $\frac{e^2}{r_{12}}$ in the potential energy. Then the energy of the system will be the sum of the energies of the two electrons separately:

$$E = E_1 + E_2. \tag{5}$$

The function ψ in (3) depends on the six co-ordinates x_1, y_1, z_1, x_2, y_2, z_2, but to the first approximation, *i.e.*, for the unperturbed system, ψ will be a product of two functions each of which depends on the co-ordinates of one electron only:

$$\psi = \psi(1)\,\psi(2). \tag{6}$$

The unperturbed equation therefore becomes:

$$(\nabla_1^2 + \nabla_2^2)\left\{\psi(1)\,\psi(2)\right\} + \frac{8\pi^2\mu}{h^2}\left\{(E_1 + E_2) + Ne^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right)\right\} \times \{\psi(1)\,\psi(2)\} = 0$$

or

$$\psi(2)\left\{\nabla_1^2\psi(1) + \frac{8\pi^2\mu}{h^2}\left(E_1 + \frac{Ne^2}{r_1}\right)\psi(1)\right\} + \psi(1)\left\{\nabla_2^2\psi(2) + \frac{8\pi^2\mu}{h^2}\left(E_2 + \frac{Ne^2}{r_2}\right)\psi(2)\right\} = 0,$$

or, dividing out by ψ (1) ψ (2) and transposing,

$$\frac{1}{\psi(1)}\left\{\nabla_1^2\psi(1)+\frac{8\pi^2\mu}{h^2}\left(E_1+\frac{Ne^2}{r_1}\right)\psi(1)\right\}$$
$$=-\frac{1}{\psi(2)}\left\{\nabla_2^2\psi(2)+\frac{8\pi^2\mu}{h^2}\left(E_2+\frac{Ne^2}{r_2}\right)\psi(2)\right\}.$$

The left-hand side of this equation depends on (x_1, y_1, z_1) only and the right-hand side on (x_2, y_2, z_2). Each of these must therefore be an absolute constant, λ say. We get then the two equations

$$\nabla_1^2\psi(1)+\frac{8\pi^2\mu}{h^2}\left(E_1-\lambda+\frac{Ne^2}{r_1}\right)\psi(1)=0, \tag{7}$$

$$\nabla_2^2\psi(2)+\frac{8\pi^2\mu}{h^2}\left(E_2+\lambda+\frac{Ne^2}{r_2}\right)\psi(2)=0. \tag{8}$$

These equations (7) and (8) are the same as that for the hydrogen-like atom, except that the eigen-value parameter is now called $E_1 - \lambda$, $E_2 + \lambda$ respectively, instead of the simple E. We integrate these two equations according to the method of § 8·5, with the provision that we take the ground state, *i.e.*, the principal quantum number $n = 1$ for (7), and $n > 1$ for (8). Then from (34) § 8·5, we have

$$\left.\begin{aligned} E_1-\lambda&=-\frac{2\pi^2\mu N^2e^4}{h^2},\\ E_2+\lambda&=-\frac{2\pi^2\mu N^2e^4}{h^2n^2}.\end{aligned}\right\} \tag{9}$$

The unperturbed wave-function is then

$$u=\psi_1(1)\,\psi_n(2), \tag{10}$$

which satisfies the unperturbed wave-equation:

$$\nabla^2u+\frac{8\pi^2\mu}{h^2}\left(E_1+E_2+\frac{Ne^2}{r_1}+\frac{Ne^2}{r_2}\right)u=0. \tag{11}$$

But the wave-function

$$v=\psi_1(2)\,\psi_n(1), \tag{12}$$

obtained by interchanging the two electrons, is also a solution of (11). Each energy value $E_1 + E_2$ is therefore two-fold, and the problem is degenerate, except when $n = 1$. Heisenberg calls it " resonance-degeneracy ".

Therefore, as shown in the theory of the last section, equation (4), § 9·5, we take a linear combination of the two functions u, v:

$$w=\alpha u+\beta v, \tag{13}$$

where the coefficients α, β are arbitrary for the present but would be determined later.

Now suppose that in the perturbed equation:

$$\nabla^2\psi + \frac{8\pi^2\mu}{h^2}\left(E + \frac{Ne^2}{r_1} + \frac{Ne^2}{r_2} - \frac{e^2}{r_{12}}\right)\psi = 0, \tag{14}$$

the perturbation term $\frac{e^2}{r_{12}}$ is considered small, and that the eigenvalue and eigen-function of (14) are given by

$$E = E_1 + E_2 + \eta, \tag{15}$$

and

$$\psi = w + \phi, \tag{16}$$

where η and ϕ are small quantities whose products with $\frac{e^2}{r_{12}}$ may be neglected.

Substituting (15) and (16) in (14), we get:

$$\nabla^2 (w + \phi) + \frac{8\pi^2\mu}{h^2}\left(E_1 + E_2 + \eta + \frac{Ne^2}{r_1} + \frac{Ne^2}{r_2} - \frac{e^2}{r_{12}}\right)(w+\phi) = 0,$$

or, neglecting small terms of the second order,

$$\nabla^2 w + \frac{8\pi^2\mu}{h^2}\left(E_1 + E_2 + \frac{Ne^2}{r_1} + \frac{Ne^2}{r_2}\right)w + \frac{8\pi^2\mu}{h^2}\left(\eta - \frac{e^2}{r_{12}}\right)w$$

$$+ \nabla^2\phi + \frac{8\pi^2\mu}{h^2}\left(E_1 + E_2 + \frac{Ne^2}{r_1} + \frac{Ne^2}{r_2}\right)\phi = 0.$$

Remembering that w satisfies (11), and writing

$$A = \frac{8\pi^2\mu}{h^2} \cdot \frac{e^2}{r_{12}}, \quad \epsilon = \frac{8\pi^2\mu}{h^2}\eta, \tag{17}$$

we get finally the equation

$$\nabla^2\phi + \frac{8\pi^2\mu}{h^2}\left(E_1 + E_2 + \frac{Ne^2}{r_1} + \frac{Ne^2}{r_2}\right)\phi = (A - \epsilon)\, w. \tag{18}$$

As usual, for the non-homogeneous equation (18) to have a solution the right-hand side must be orthogonal to the eigen-values u and v of the homogeneous equation (11). Thus we get:

$$\int (A - \epsilon)\, w\, u\, d\tau = 0, \quad \int (A - \epsilon)\, w\, v\, d\tau = 0, \tag{19}$$

where

$$d\tau = (dx_1\, dy_1\, dz_1)\,(dx_2\, dy_2\, dz_2) = d\tau_1\, d\tau_2.$$

Substituting (13) in (19), we get the two equations:

$$\left.\begin{aligned} \alpha \int (\mathrm{A} - \epsilon)\, u^2 d\tau + \beta \int (\mathrm{A} - \epsilon)\, uv\, d\tau = 0 \\ \alpha \int (\mathrm{A} - \epsilon)\, uv d\tau + \beta \int (\mathrm{A} - \epsilon)\, v^2\, d\tau = 0 \end{aligned}\right\}. \tag{20}$$

Suppose u and v are normalised to 1, so that

$$\int u^2 d\tau = 1, \quad \int v^2 d\tau = 1. \tag{21}$$

Further, we have

$$\begin{aligned} \int uv d\tau &= \int \psi_1 (1)\ \psi_n (1)\, d\tau_1 . \int \psi_1 (2)\ \psi_n (2)\, d\tau_2 \\ &= 0. \end{aligned} \tag{22}$$

We write

$$a_{11} = \int \mathrm{A}\ u^2\ d\tau, \quad a_{12} = a_{21} = \int \mathrm{A}\ uv d\tau, \quad a_{22} = \int \mathrm{A} v^2 d\tau. \tag{23}$$

Substituting (21), (22) and (23) in (20), we get

$$\left.\begin{aligned} \alpha\, (a_{11} - \epsilon) + \beta\, a_{12} = 0, \\ \alpha\, a_{12} + \beta\, (a_{22} - \epsilon) = 0. \end{aligned}\right\} \tag{24}$$

Eliminating α and β, we get a quadratic equation for ϵ:

$$(a_{11} - \epsilon)\, (a_{22} - \epsilon) - a_{12}{}^2 = 0. \tag{25}$$

But $\mathrm{A} = \frac{8\pi^2\mu}{h^2}\frac{e^2}{r_{12}}$ is symmetrical in both the electrons, v arises from u by the interchange of both electrons, and the integration with respect to $d\tau$ considers both the electrons equally, therefore we have

$$a_{11} = a_{22}. \tag{26}$$

Hence, (25) becomes

$$(a_{11} - \epsilon)^2 - a_{22}{}^2 = 0. \tag{27}$$

The two roots of the equation (27) are therefore,

$$\epsilon_1 = a_{11} - a_{12}, \ \epsilon_2 = a_{11} + a_{12}. \tag{28}$$

The eigen-value $\mathrm{E}_1 + \mathrm{E}_2$ of the unperturbed problem (11) breaks up into the two somewhat different values

$$\mathrm{E}_1 + \mathrm{E}_2 + \epsilon_1 \frac{h^2}{8\pi^2\mu}, \quad \mathrm{E}_1 + \mathrm{E}_2 + \epsilon_2 \frac{h^2}{8\pi^2\mu}. \tag{29}$$

Substituting (28) in (24) we have for ϵ_1:

$$\alpha\, (a_{11} - \epsilon_1) + \beta\, a_{12} = 0,$$

or

$$\alpha\, a_{12} + \beta\, a_{12} = 0,$$

or

$$\alpha + \beta = 0, \beta = -\alpha. \tag{30}$$

Substituting (30) in (13) we get for the eigen-function corresponding to ϵ_1 :

$$w_1 = au - av = a(u - v). \tag{31}$$

Similarly, from (24) and (28), we have for ϵ_2 :

$$a(a_{11} - \epsilon_2) + \beta a_{12} = 0$$

or

$$-aa_{12} + \beta a_{12} = 0$$

or

$$-a + \beta = 0, \text{ or } \beta = a.$$

Then from (13) we get for the second eigen-function,

$$w_2 = au + av = a(u + v). \tag{32}$$

We have finally to determine the value of a. For this we use the condition that both w_1 and w_2 should be normalised to 1 :

$$\int w_1^2 d\tau = 1, \tag{33}$$

i.e.,

$$a^2 \int (u - v)^2 d\tau = 1,$$

which on account of (21) and (22) reduces to

$$2a^2 = 1, \quad a = \frac{1}{\sqrt{2}}. \tag{34}$$

Thus to a first approximations, the two eigen-functions are

$$w_1 = \frac{1}{\sqrt{2}}(u - v) = \frac{1}{\sqrt{2}}\left\{\psi_1(1)\,\psi_n(2) - \psi_1(2)\,\psi_n(1)\right\}, \tag{35}$$

$$w_2 = \frac{1}{\sqrt{2}}(u + v) = \frac{1}{\sqrt{2}}\left\{\psi_1(1)\,\psi_n(2) + \psi_1(2)\,\psi_n(1)\right\}. \tag{36}$$

We see that one of these eigen-functions, *viz.*, w_1 is anti-symmetrical in the two electrons, *i.e.*, if we interchange the two electrons, its sign is changed. The other function, w_2, is symmetrical in the two electrons, *i.e.*, it does not change if the two electrons are interchanged.

We see therefore that the neutral helium atom has two distinct sets of eigen-functions, (*a*) anti-symmetrical and (*b*) symmetrical. We know also that the eigen-functions characterise in some way the stationary states of the atom. We shall call the symmetric eigen-functions the Para-terms, or simply the P-terms, and the anti-symmetric ones Ortho-terms, or the O-terms.

We shall prove now that the two systems of terms (35) and (36) do not combine with each other, *i.e.*, that there can be no transition between a P-term and an O-term.

Let f (1, 2) be the matrix representing the radiation from the atom. It is of course symmetrical in both the electrons 1 and 2, *i.e.*,

$$f(2, 1) = f(1, 2). \tag{37}$$

If the element $f_{m,n}$ (1, 2) vanishes for every component of f in every direction, this means that the transition $m \longrightarrow n$ is forbidden. For a transition from the state $\frac{1}{\sqrt{2}}$ $(u + v)_m$ to the state $\frac{1}{\sqrt{2}}$ $(u - v)_n$, we get

$$f_{mn} = \tfrac{1}{2} \iint f(1, 2)\,(u + v)_m\,(u - v)_n\,d\tau$$

or

$$f_{mn} = \tfrac{1}{2} \iint f(1, 2)\,(u + v)_m\,(u - v)_n\,d\tau_1\,d\tau_2. \tag{38}$$

On account of the anti-symmetrical character of $u - v$, the integrand in (38) changes sign if we interchange the two electrons, whereas f_{mn} being symmetrical, is not affected by such an interchange, *i.e.*,

$$f_{mn} = -\tfrac{1}{2} \iint f(1,2)\,(u + v)_m\,(u - v)_n\,d\tau_1\,d\tau_2. \tag{39}$$

Comparing (38) and (39) we find

$$f_{mn} = 0 \tag{40}$$

for all m and all n, showing that there is no transition from a P-term to an O-term. The existence of the two non-combining sets of terms, known as the Para-Helium and Ortho-Helium lines, has been abundantly proved experimentally, and it was one of the problems which the classical theory and the older quantum theory of Bohr could not explain.

9·7. *Systems Containing Similar Particles.*

After the first work of Heisenberg on the neutral helium atom containing only two electrons, and on the lines suggested by him, many workers such as Dirac, Hund, Wigner, Heitler and others, attacked the problem of many bodies involving several similar particles. A detailed treatment of this would require a knowledge of the theory of permutation-groups. The permutation-groups come into play on account of the fundamental fact, as mentioned in the last section, that particles identical in their nature cannot be distinguished from one another by their position or motion, so that if we make any permutation, or interchange, of any number of the particles among themselves, the subsequent state is absolutely indistinguishable from the original state as far as observation is concerned.

We shall content ourselves with giving only the conclusions arrived at by a use of the theory of groups. The following account is taken from Dirac's work.

The results given in the last section for only two electrons are found to hold in the general case. It is found that the eigen-functions are divided into a finite number of sets as regards their symmetry character. These sets are mutually non-combining or exclusive, that is to say, that there is no transition between a member of one set and a member of another set. Transitions can take place only between two members of the same set.

It is further possible to deduce the following theorems about the symmetry-character of the eigen-functions.

THEOREM 1. If an eigen-function is initially symmetrical, it remains always symmetrical.

THEOREM 2. If an eigen-function is anti-symmetrical, it remains always anti-symmetrical.

THEOREM 3. The property of an eigen-function being symmetrical or anti-symmetrical remains invariant under a canonical transformation.

These theorems show that the property of being symmetrical or anti-symmetrical is a property of the states themselves, and not merely a property of the eigen-functions. Thus we can say that the states themselves are symmetrical or anti-symmetrical.

Moreover, the invariance and permanance of the symmetry character of the states means that for some particular kind of particles it is quite possible for only symmetrical or only anti-symmetrical states to occur in nature. The question is decided in the case of photons by making use of Planck's law of radiation, and in the case of electrons by making use of Pauli's exclusion principle.

It is found that only when we assume the symmetric states for photons, do we get a statistical mechanics leading to Planck's law of black-body radiation. This statistics for the photons has been developed by S. N. Bose and A. Einstein.

For the electrons we use Pauli's exclusion principle, which has been inferred from general experimental evidence on atomic structure, and is universally accepted. This principle, as we know, postulates that, speaking in terms of Bohr's theory, no two electrons can ever be in the same orbit. If we wish to avoid the conception of the electron orbits, we can put Pauli's exclusion principle in the form that each individual state can be associated with only one electron.

Now Dirac remarks that if Pauli's exclusion principle is to hold for a system of electrons, such a system must be described by an anti-symmetrical eigen-function. For example, consider a system of two electrons. As found in (35) § 9·6, an antisymmetrical eigen-function for these two electrons is

$$w = \frac{1}{\sqrt{2}} \left\{ \psi_m(1)\ \psi_n(2) - \psi_m(2)\ \psi_n(1) \right\}. \tag{1}$$

In this function if for ψ_m (2), ψ_n (2), we write ψ_m (1), ψ_n (1) respectively, *i.e.*, bring the second electron in the orbit of the first electron, then w vanishes.

Thus, if we assume that for electrons only anti-symmetrical states occur, then we get the result that there are no states with two electrons in the same orbit, which is just Pauli's exclusion principle. The statistics for electrons has been developed by Dirac and E. Fermi, and has been applied to the electron theory of conduction in metals by Sommerfeld and others.

9·8. *Perturbation Theory in Matrix Mechanics.*

In § 9·3 and § 9·5 we gave the wave-mechanical perturbation theory. We remarked there that if the Hamiltonian of a dynamical system depends on a parameter ϵ, *i.e.*, $H = H(p, q, \epsilon)$, and if the eigen-values are known for the particular value $\epsilon = 0$, then for small values of ϵ, the perturbation in the eigen-values can be determined by an expansion in series :

$$H(p, q, \epsilon) = H_0(p, q) + \epsilon H_1(p, q) + \epsilon^2 H_2(p, q) + \cdots \tag{1}$$

We shall now give an account of the perturbation theory as developed in matrix mechanics. It is found in applications, that often the matrix method gives the result more quickly and in a more elegant form.

The problem is as follows. The Hamiltonian function $H = H_0(p, q)$ defines a certain dynamical system, and the solution of the problem

$$pq - qp = \frac{h}{2\pi i}\,\delta, \quad H_0(p, q) = E_0 = \text{diagonal matrix}, \tag{2}$$

is already determined, that is to say that we know certain matrices p_0, q_0, which actually satisfy the equations (2). We have now a slightly different problem

$$pq - qp = \frac{h}{2\pi i}\,\delta, \quad H(p, q) = E = \text{diagonal matrix}, \tag{3}$$

where $H(p, q)$ is supposed to be capable of expansion in the series (1), and consequently E can be expanded in the series

$$E = E_0 + \epsilon E_1 + \epsilon^2 E_2 + \cdots \tag{4}$$

We try to build up the solution of (3) with the help of the solution of the problem (2), given by the matrices p_0, q_0. We saw in § 6·9 (4) that all transformation matrices in quantum mechanics must be unitary. Let u be such a matrix with the property

$$u^\dagger = u^{-1}, \tag{5}$$

then the matrices p_0, q_0 would be transformed to

$$p = u^{-1} p_0 u, \quad q = u^{-1} q_0 u. \tag{6}$$

We know from Theorem 1 § 6·9 (1) that p, q are hermitian, and from § 6·9 (4) that they satisfy the quantum condition:

$$pq - qp = \frac{h}{2\pi i}\,\delta. \tag{7}$$

Further, we get from Theorem 2 § 6·9 (2),

$$H(p, q) = u^{-1} H(p_0, q_0)\, u, \tag{8}$$

or, on account of (3)

$$u^{-1} H(p_0, q_0)\, u = E,$$

i.e.,

$$H(p_0, q_0)\, u = u E, \tag{9}$$

Corresponding to the expansions (1) and (4), we set

$$u = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \cdots, \tag{10}$$

and get from (9), on account of (1) and (4):

$$\{H_0(p_0, q_0) + \epsilon H_1(p_0, q_0) + \epsilon^2 H_2(p_0, q_0) + \cdots\} \times \{u_0 + \epsilon u_1 + \epsilon^2 u_2 + \cdots\}$$
$$= \{u_0 + \epsilon u_1 + \epsilon^2 u_2 + \cdots\}\{E_0 + \epsilon E_1 + \epsilon^2 E_2 + \cdots\}. \tag{11}$$

Equating the coefficients of equal powers of ϵ, and taking account of the relation $H_0\,(p_0, q_0) = E_0$ given by (2), we obtain

$$E_0\,u_0 = u_0\,E_0, \tag{12}$$

$$E_0\,u_1 + H_1\,u_0 = u_0\,E_1 + u_1\,E_0, \tag{13}$$

$$E_0\,u_2 + H_1\,u_1 + H_2\,u_0 = u_0\,E_2 + u_1\,E_1 + u_2\,E_0. \tag{14}$$

When higher approximations are required, we can equate coefficients of ϵ^3, ϵ^4, etc.

Undisturbed System Non-degenerate.—In this case, no two energy levels are equal, *i.e.*,

$$E_0\,(nn) \neq E_0\,(mm) \text{ if } n \neq m, \tag{15}$$

so that all the terms of the diagonal matrix E_0 are different. From (12) we obtain, therefore,

$$\sum_k \{E_0\,(mk)\,u_0\,(kn) - u_0\,(mk)\,E_0\,(kn)\} = 0,$$

i.e.,

$$u_0\,(mn)\,\{E_0\,(mm) - E_0\,(nn)\} = 0,$$

i.e.,

$$hu_0\,(mn)\,\nu_0\,(mn) = 0. \tag{16}$$

This gives, since $\nu_0\,(mn) \neq 0$ if $m \neq n$,

$$u_0\,(mn) = 0 \text{ if } m \neq n, \tag{17}$$

showing that u_0 is a diagonal matrix. But the diagonal terms cannot be determined from (16). To determine these, we have to use the fact that u is a unitary matrix. From (5) we get then $uu^\dagger = \delta$, *i.e.*,

$$(u_0 + \epsilon\,u_1 + \epsilon^2\,u_2 + \cdots)\,(u_0^\dagger + \epsilon\,u_1^\dagger + \epsilon^2\,u_2 + \cdots) = \delta\;;$$

hence

$$u_0\,u_0^\dagger = \delta, \text{ and } u_0\,u_1^\dagger + u_1\,u_0^\dagger = 0.$$

Therefore, we have, since u_0 is a diagonal matrix,

$$u_0\,(nn)\,u_0^\dagger\,(nn) = 1, \textit{ i.e.}, \quad |\,u_0\,(nn)\,|^2 = 1. \tag{18}$$

Thus the phase of $u_0\,(nn)$ is indeterminate, and if we write

$$u_0\,(nn) = |\,u_0\,(nn)\,|\,e^{i\phi(nn)} = e^{i\phi(nn)},$$

we can absorb $e^{i\phi(nn)}$ into the usual unwritten time factor, and write $u_0\,(nn) = 1$. Thus we have effectively

$$u_0 = \delta \text{ (unit matrix)}. \tag{19}$$

The equation (13) now becomes

$$E_0\,u_1 + H_1 = E_1 + u_1\,E_0, \text{ or } E_0\,u_1 - u_1\,E_0 = E_1 - H_1. \tag{20}$$

Taking the mn component of both sides, we obtain:

$$\sum_k \{E_0(mk)\, u_1(kn) - u_1(mk)\, E_0(kn)\} = E_1(mn) - H_1(mn),$$

or

$$u_1(mn)\{E_0(mm) - E_0(nn)\} = E_1(mn) - H_1(mn),$$

or

$$h\, u_1(mn)\, \nu_0(mn) = E_1(mn) - H_1(mn). \tag{21}$$

But E_1 is a diagonal matrix, so that if $m \neq n$, then

$$h\, u_1(mn)\, \nu_0(mn) = -H_1(mn), \; (m \neq n), \tag{22}$$

and if $m = n$, $\nu_0(mn) = 0$, so that

$$E_1(nn) = H_1(nn). \tag{23}$$

Thus we see that H_1 is not a diagonal matrix, but its diagonal terms are those of E_1.

The "time mean" $\bar{a}$ of a matrix a, is defined, in quantum mechanics, as the matrix formed by writing all the terms of a zero, except the constant terms, which for a general matrix are only the diagonal terms:

$$\bar{a} = (a_{mn}\, \delta_{mn}). \tag{24}$$

From (23) we obtain therefore

$$E_1 = \bar{H}_1. \tag{25}$$

Thus, we get the eigen-value

$$E = E_0 + \epsilon \bar{H}_1, \tag{26}$$

correct to the first order of approximations.

This result is analogous to that of the classical theory, and also to that found in the wave-mechanics.

We shall now calculate the second approximation.

From (14) we obtain, since $u_0 = \delta$,

$$E_0 u_2 - u_2 E_0 = E_2 - H_2 + u_1 E_1 - H_1 u_1. \tag{27}$$

Therefore, on taking the mn component we get

$$\sum_k \{E_0(mk)\, u_2(kn) - u_2(mk)\, E_0(kn)\} = E_2(mn) - H_2(mn) + \sum_k \{u_1(mk)\, E_1(kn) - H_1(mk)\, u_1(kn)\},$$

or,

$$u_2(mn)\{E_0(mm) - E_0(nn)\} = E_2(mn) - H_2(mn) + u_1(mn)\, E_1(nn) - \sum_k \{H_1(mk)\, u_1(kn)\},$$

or

$$h\,u_2\,(mn)\,\nu_0\,(mn) = \mathrm{E}_2\,(mn) - \mathrm{H}_2\,(mn) + u_1\,(mn)\,\mathrm{E}_1\,(nn) - \sum_k \{\mathrm{H}_1\,(mk)\,u_1\,(kn)\}. \tag{28}$$

Writing $m = n$ in (28), we find, since $\nu_0\,(nn) = 0$,

$$\mathrm{E}_2\,(nn) = \sum_k \mathrm{H}_1\,(nk)\,u_1\,(kn) + \mathrm{H}_2\,(nn) - u_1\,(nn)\,\mathrm{E}_1\,(nn). \tag{29}$$

But, from the condition $uu^\dagger = \delta$, we got as the first approximation the result (19), *viz.*, $u_0 = \delta$. Substituting this in the second approximation $u_0\,u_1{}^\dagger + u_1\,u_0{}^\dagger = 0$, we obtain

$$u_1 + u_1{}^\dagger = 0. \tag{30}$$

Taking the nn component, we find

$$u_1\,(nn) + u_1{}^\dagger\,(nn) = 0.$$

Writing

$$u_1\,(nn) = c\,e^{i\phi(nn)},$$

with c real and the phase ϕ arbitrary, we get

$$c\,\{e^{i\phi(nn)} + e^{-i\phi\,(nn)}\} = 0,$$

or

$$2c\cos\phi\,(nn) = 0.$$

Since the phase ϕ is arbitrary, we conclude that $c = 0$, and therefore

$$u_1\,(nn) = 0. \tag{31}$$

Substituting the values of $u_1\,(mn)$ and $u_1\,(nn)$ from (22) and (31) in (29), we obtain

$$\begin{aligned}\mathrm{E}_2\,(nn) &= \mathrm{H}_2\,(nn) - \sum_k{}' \frac{\mathrm{H}_1\,(nk)\,\mathrm{H}_1\,(kn)}{h\nu_0\,(kn)}\\ &= \mathrm{H}_2\,(nn) + \sum_k{}' \frac{\mathrm{H}_1\,(nk)\,\mathrm{H}_1\,(kn)}{h\nu_0\,(nk)},\end{aligned} \tag{32}$$

where $\sum_k{}'$ means that the term $k = n$ has been omitted from the summation. The disturbed eigen-value is thus found to the second approximation :

$$\mathrm{E} = \mathrm{E}_0 + \epsilon\,\bar{\mathrm{H}} + \epsilon^2 \left\{\mathrm{H}_2\,(nn) + \sum_k{}' \frac{\mathrm{H}_1\,(nk)\,\mathrm{H}_1\,(kn)}{h\nu_0\,(nk)}\right\}. \tag{33}$$

Undisturbed System Degenerate.—In this case r of the discrete energy-levels are equal, *i.e.*,

$$\mathrm{E}_0\,(n_1n_1) = \mathrm{E}_0\,(n_2n_2) = \cdots = \mathrm{E}_0\,(n_rn_r). \tag{34}$$

Then from (16), *viz.*, $h\nu_0\,(mn)\,u_0\,(mn) = 0$, we conclude, as before, that $u_0\,(nn)$, *i.e.*, the diagonal terms of the matrix u_0, are not zero. But in the case of degeneracy, we find that in addition to the diagonal terms, some of the non-diagonal elements are also not zero, *viz.*, those terms $u_0\,(mn)$ for which m, n are any pair of the set $n_1, n_2, \cdots, n_r$. Other non-diagonal terms for which m, n are not a pair of this set, vanish.

Now since $\nu_0\,(mn) = 0$ when $m = n$, or when m, n are any pair of the set $n_1, n_2, \cdots, n_r$, we see that the matrix u_0 consists of constant terms only, though it is not a diagonal matrix.

From the condition $u_0\,u_0{}^\dagger = \delta$, we get

$$u_0\,(nn)\,u_0{}^\dagger\,(nn) = 1 \text{ if } n \neq n_1, n_2, \cdots, n_r\,;$$

i.e.,

$$|\,u_0\,(nn)\,|^2 = 1, \;\; u_0\,(nn) = 1 \text{ if } n \neq n_1, n_2, \cdots, n_r,$$

and

$$u_0\,(n_k, n_k) \neq 1 \text{ if } k = 1, 2, \cdots, r.$$

The equation (13) then becomes

$$E_0\,u_1 - u_1\,E_0 = u_0\,E_1 - H_1\,u_0.$$

Taking the mn component, we obtain

$$\sum_k \{E_0\,(mk)\,u_1\,(kn) - u_1\,(mk)\,E_0\,kn)\}$$
$$= \sum_k \{(u_0\,(mk)\,E_1\,(kn) - H_1\,(mk)\,u_0\,(kn)\},$$

or

$$u_1\,(mn)\,\{E_0\,(mm) - E_0\,(nn)\} = u_0\,(mn)\,E_1\,(nn) - \sum_k H_1\,(mk)\;u_0\,(kn). \tag{35}$$

Thus, if $m \neq n$ and $m, n \neq n_1, n_2, \cdots, n_r$, then since $u_0\,(nn) = 1$, we get

$$hu_1\,(mn)\,\nu_0\,(mn) = -\,H_1\,(mn)\,u_0\,(nn) = -\,H_1\,(mn), \tag{36}$$

which shows that $H_1\,(mn)$ has a value not zero.

Writing $m = n \neq n_1, n_2, \cdots, n_r$, in (35) we get

$$0 = E_1\,(nn) - H_1\,(nn)\;, \text{ or } E_1\,(nn) = H_1\,(nn). \tag{37}$$

Thus the diagonal elements of H_1 still give those of E_1, except when $n = n_1, n_2, \cdots, n_r$.

To find the remaining r diagonal elements $E_1\,(n_1\,n_1)$, $E_1\,(n_2 n_2), \cdots, E_1\,(n_r n_r)$, we put $n = n_1$ in (35), and get

$$hu_1\,(mn_1)\,\nu_0\,(mn_1) = u_0\,(mn_1)\,E_1\,(n_1 n_1) - \{H_1\,(mn_1)\,u_0\,(n_1 n_1)$$
$$+ H_1\,(mn_2)\,u_0\,(n_2 n_1) + \cdots + H_1\,(mn_r)\,u_0\,(n_r n_1)\}. \tag{38}$$

We put m successively equal to $n_1, n_2, \cdots, n_r$ in (38), and note that $\nu_0\,(n_k n_l) = 0$ for $k, l = 1, 2, \cdots, r$. Then we obtain

$$u_0\,(n_k n_1)\, E_1\,(n_1 n_1) - \{H_1\,(n_k n_1)\, u_0\,(n_1 n_1) + H_1\,(n_k n_2)\, u_0\,(n_2 n_1) + \cdots + H_1\,(n_k n_r)\, u_0\,(n_r n_1)\} = 0,$$

$$(k = 1, 2, \cdots, r). \tag{39}$$

This gives the following secular equation for $E_1\,(n_1 n_1)$, provided we eliminate $u_0\,(n_1 n_1),\ u_0\,(n_2 n_1),\ \cdots,\ u_0\,(n_r, n_1)$:

$$\begin{vmatrix} H_1\,(n_1 n_1) - E_1\,(n_1 n_1), & H_1\,(n_1 n_2), & \cdots\cdots, & H_1\,(n_1 n_r), \\ H_1\,(n_2 n_1), & H_1\,(n_2 n_2) - E_1\,(n_1 n_1), & \cdots, & H_1\,(n_2 n_r), \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ H_1\,(n_r n_1), & \cdots\cdots & \cdots\cdots, & H_1\,(n_r n_r) - E_1\,(n_1 n_1). \end{vmatrix} = 0 \tag{40}$$

If we had put $n = n_2, \cdots, n_r$ instead of n_1 in (35), we would have obtained the same equation (40), except that $E_1\ (n_1 n_1)$ would have been replaced by $E_1\,(n_2 n_2), \cdots, E_1\,(n_r n_r)$. Thus $E_1\,(n_1 n_1),\ E_1\,(n_2 n_2), \cdots E_1\,(n_r n_r)$ are the r roots of the secular equation

$$\begin{vmatrix} H_1\,(n_1 n_1) - \lambda, & H_1\,(n_1 n_2), & \cdots\cdots, & H_1\,(n_1 n_r) \\ H_1\,(n_2 n_1), & H_1\,(n_2 n_2) - \lambda, & \cdots, & H_1\,(n_2 n_r) \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ H_1\,(n_r n_1), & \cdots\cdots & \cdots\cdots, & H_1\,(n_r n_r) - \lambda. \end{vmatrix} = 0 \tag{41}$$

The r roots of the equation (41) give the remaining elements of the diagonal matrix E_1, showing that the r coincident levels H_0 split up into r different levels.

Now, since $\nu_0\,(mn) = 0$, not only when $m = n$, but also when m, n are any pair of the set $n_1, \cdots, n_r$, we see that the matrix H_1 has constant elements not only in the leading diagonal, but also at the crossing points of the n_1th, $\cdots$, n_rth rows with the n_1th, $\cdots$, n_rth columns. Then, from the definition of the time mean of a matrix, we know that $\bar{H}_1$ is not a diagonal matrix for the degenerate case, since, some non-diagonal terms of H_1, *viz.*, $H_1\,(n_k, n_l)$ where $k \neq l$ and and $k, l = 1, 2, \cdots r$, are also constant, and cannot be made zero when forming the time mean. Then we transform the matrix H_1 so that it becomes the diagonal matrix E_1, thus

$$\bar{S}^1 H_1 S = E_1, \text{ or } \bar{H}_1 S = S E_1. \tag{42}$$

Taking the $n_k\, m$ element, this gives

$$\sum_l \mathrm{H}_1\,(n_k l)\,\mathrm{S}\,(l\, m) = \sum_l \mathrm{S}\,(n_k l)\,\mathrm{E}_1\,(lm),$$

or since H_1 has only terms $\mathrm{H}_1\,(nn)$ and $\mathrm{H}_1\,(n_k n_l)$, where $k, l = 1, 2, \cdots, r$, we obtain

$$\sum_{l=1}^{r} \mathrm{H}_1\,(n_k n_l)\,\mathrm{S}\,(n_l m) = \mathrm{S}\,(n_k m)\,\mathrm{E}_1\,(mm), \qquad (k, m = 1, 2, \cdots, r). \qquad (43)$$

On eliminating the S's from this system, we obtain the same equation for the determination of the r terms of E_1 as (41).

This shows that if the undisturbed system is degenerate, then $\bar{\mathrm{H}}_1$ is a matrix of constant terms, though not a diagonal matrix. If a transformation $\bar{\mathrm{S}}^1\,\bar{\mathrm{H}}_1\,\mathrm{S}$ is carried out so that this becomes a diagonal matrix E_1, then the eigen-values of the perturbed system are, to a first approximation,

$$\mathrm{E} = \mathrm{E}_0 + \epsilon\,\mathrm{E}_1. \qquad (44)$$

The process can be carried to second and higher approximations, but we shall not enter into details here.

CHAPTER X.

PHYSICAL INTERPRETATIONS AND HEISENBERG'S PRINCIPLE OF INDETERMINACY.

10 · 1 (1). *Physical Significance of the Wave-Function. Schrödinger's Interpretation.*

So far we have been using the function ψ freely without stopping to consider what it actually signifies. The important fact for us was that the ψ-equation gave correct energy values for any atomic system.

It was Schrödinger again who first gave a physical interpretation of the wave-function. To set forth his arguments, we consider the problem of a single electron with mass m and charge ϵ. Using Cartesian co-ordinates we have the following wave-equation for the electron :

$$\frac{1}{m}\nabla^2\psi + \frac{8\pi^2}{h^2}(E - V)\psi = 0. \tag{1}$$

Multiplying this equation by the factor $e^{-2\pi i\frac{E}{h}t}$, we get

$$\frac{1}{m}\nabla^2\left(\psi e^{-2\pi i\frac{E}{h}t}\right) + \frac{8\pi^2}{h^2}(E - V)\left(\psi e^{-2\pi i\frac{E}{h}t}\right) = 0,$$

or, writing

$$\phi(x, y, z\,;\,t) = \psi(x, y, z)\,e^{-2\pi i\frac{E}{h}t}, \tag{2}$$

we get for the wave-equation :

$$\frac{1}{m}\nabla^2\phi + \frac{8\pi^2}{h^2}(E - V)\phi = 0. \tag{3}$$

From equations (1), (2) of § 7 · 5, we know that the true wave-function is ϕ, and that it depends on time according to the law (2). Moreover, the equation (2) gives us also

$$\frac{\partial\phi}{\partial t} = -\frac{2\pi i}{h}E\phi,$$

so that in agreement with (6) § 8 · 6 (1), we find :

$$E\phi = -\frac{h}{2\pi i}\frac{\partial\phi}{\partial t}. \tag{4}$$

Substituting (4) in (3) we get

$$\frac{1}{m}\nabla^2\phi - \frac{8\pi^2}{h^2}V\phi + \frac{4\pi i}{h}\frac{\partial\phi}{\partial t} = 0. \tag{5}$$

Similarly, if ϕ^* is the complex conjugate of ϕ, it must satisfy the equation

$$\frac{1}{m}\nabla^2\phi^* - \frac{8\pi^2}{h^2}V\phi^* - \frac{4\pi i}{h}\frac{\partial\phi^*}{\partial t} = 0. \tag{6}$$

Multiplying (5) by ϕ^* and (6) by ϕ, and subtracting, we get:

$$\frac{1}{m}(\phi^*\nabla^2\phi - \phi\nabla^2\phi^*) + \frac{4\pi i}{h}\frac{\partial}{\partial t}(\phi\phi^*) = 0. \tag{7}$$

Multiplying (7) by ϵ and rearranging, we have:

$$\frac{\partial}{\partial t}(\epsilon\phi\phi^*) + \frac{\epsilon h}{4\pi i m}(\phi^*\nabla^2\phi - \phi\nabla^2\phi^*) = 0. \tag{8}$$

Let us define four functions S_x, S_y, S_z and ρ by the following equations:

$$\begin{aligned} \rho &= \epsilon\phi^*\phi, \\ S_x &= \frac{\epsilon h}{4\pi i m}\left(\phi^*\frac{\partial\phi}{\partial x} - \phi\frac{\partial\phi^*}{\partial x}\right), \\ S_y &= \frac{\epsilon h}{4\pi i m}\left(\phi^*\frac{\partial\phi}{\partial y} - \phi\frac{\partial\phi^*}{\partial y}\right), \\ S_z &= \frac{\epsilon h}{4\pi i m}\left(\phi^*\frac{\partial\phi}{\partial z} - \phi\frac{\partial\phi^*}{\partial z}\right). \end{aligned} \tag{9}$$

Then we see immediately that (8) reduces to

$$\frac{\partial\rho}{\partial t} + \frac{\partial S_x}{\partial x} + \frac{\partial S_y}{\partial y} + \frac{\partial S_z}{\partial z} = 0. \tag{10}$$

If we interpret ρ to represent the density of electric charge, and if S_x, S_y, S_z denote the components of current density, then equation (10) is nothing but the equation of continuity in the theory of electric currents. Obviously, all the four expressions in (9) give us real quantities. According to Schrödinger, therefore, the density of electric charge at the point (x, y, z) is proportional to $\phi^*\phi = \psi^*\psi$. But this view of Schrödinger's was found to be untenable when applied to the general case of a system more complicated than a single electron. In fact, Schrödinger himself had attempted previously to explain the relation between waves and corpuscles by suggesting that the electron is constituted by a group of waves of neighbouring frequencies—in other words, by a " wave-packet ". But it was soon pointed out that this simple and interesting view could not be seriously considered. It was open to many objections. Just as a wave crest produced at any point of an otherwise smooth surface of water spreads out and disappears before long, so would an electron wave-packet be in general very soon dissipated. Further, when the electron is diffracted by a crystal in the experiments of G. P. Thomson and others, the wave-packet would be completely dispersed and destroyed, showing that the particles would have no stable existence.

Another hypothesis, put forward by L. de Broglie, was to consider the particle merely as a singularity in a wave-phenomenon. As we know from the theory of functions, a singularity of a function is that point in the domain of the variables where the function becomes discontinuous. For the special case of the uniform motion of a single electron, de Broglie found a solution of the wave-equation which contained a singularity moving with the velocity of the particle. But in the general case of a non-uniform motion of the electron, the existence of such a moving singularity could not be proved, and de Broglie's hypothesis had to be dropped.

10 · 1 (2). *Born's Interpretation.*

These hypotheses of Schrödinger and de Broglie were soon replaced by a suggestion first put forward by Max Born, and now accepted generally.

Born was led to the formulation of this point of view by a remark of Einstein about the relation between the wave-field and the light-quanta. Einstein had said that the waves are there only for the purpose of showing the way, so to speak, to the light quanta. In this sense, he spoke of the waves being a sort of a "ghost-field" which determines the probability for the light-quantum to go on a definite way.

This idea of Einstein's was taken over by Born for the material waves also, with this difference, however, that the wave-field is not just a mathematical fiction, but possesses as much reality as the corpuscle. Born remarked further that the quantum mechanics so far developed solves problems connected with the stationary states only, whereas it should actually be capable of formulating and solving the transition phenomena as well.

Born adopted the following point of view, which is now gaining universal recognition and causing a revolution in our whole scientific outlook. He pointed out that the new mechanics does not answer the question : " How does a particle move ?" ; rather it answers only the question : " What is the probability that a particle should move in a given way ?" Born assumed, therefore, that the particles are always accompanied by a wave-process. These are the well-known de Broglie-Schrödinger phase waves we have considered above. The square of the amplitude of these waves determines the probability for a particle to be present in a state of motion corresponding to the given phase wave. The underlying idea has been very clearly stated by Born : " According to this view, the whole course of events is determined by the laws of probability, which is given by the de Broglie wave associated with the state. A mechanical process is accompanied by a wave-process, the guiding wave, described by Schrödinger's equation, the significance of which is that it gives the probability of a definite course of the machanical process. If, for example, the amplitude of the guiding wave is zero at a certain point in space, this means that the probability of finding the electron at this point is vanishingly small." (*Atomic Physics*, p. 83.)

Born's interpretation is thus a statistical one, in which he assumes that the function ψ is not a characteristic of the individual atom, but determines the average behaviour of the whole system.

To take a simple case, we consider an atom with many electrons. Then from (5) § 10 · 1 (1), we have the Schrödinger equation for such an atom :

$$\nabla^2\phi - \frac{8\pi^2\mu}{h^2} V\phi + \frac{4\pi i\mu}{h}\frac{\partial\phi}{\partial t} = 0, \tag{1}$$

where we have written μ for the mass of the electron, and of which, on account of (2) § 10 · 1 (1), we have a solution

$$\phi_n\,(x, y, z\,;\,t) = \psi_n\,(x, y, z)\,e^{\frac{-2\pi i}{h}E_n t}. \tag{2}$$

Here E_n is the eigen-value and ψ_n the eigen-function belonging to the nth stationary state.

The most general solution of (1) is found by superposing the solutions (2), so that we write :

$$\psi\,(x, y, z\,;\,t) = \sum_n c_n\psi_n\,(x, y, z)\,e^{\frac{-2\pi i}{h}E_n t}, \tag{3}$$

where the c_n are the coefficients of the "Fourier" expansion. Since the eigen-functions ψ_n are mutually orthogonal and normalised, we get just as in (10) § 8 · 6 (2):

$$c_n e^{\frac{-2\pi i}{h} E_n t} = \iiint \psi_n^* (x, y, z) \psi (x, y, z ; t) dv, \tag{4}$$

where dv is the volume element $dx\, dy\, dz$. Therefore on taking the square of the modulus, we get:

$$| c_n |^2 = \left| \iiint \psi_n^* (x, y, z) \psi (x, y, z ; t) dv \right|^2 . \tag{5}$$

Born postulates that $| c_n |^2$ is the probability that the atom is in the nth state.

Now suppose that when the atom is in this nth state, it is acted upon by a perturbing force which lasts a finite time, and that under the influence of this perturbation the atom makes a transition to another state. Born has proved that in this case the Fourier coefficients c_n undergo an orthogonal transformation:

$$C_n = \sum_m b_{mn} c_m, \tag{6}$$

where, on account of the orthogonality,

$$\begin{aligned} \sum_k b_{mk} b_{nk}^* &= 1 \text{ when } m = n, \\ &= 0 \text{ when } m \neq n. \end{aligned} \tag{7}$$

Born now puts forward the supplementary hypothesis that $| b_{nm} |^2$ represents the probability for the atom to make a transition from the state n to the state m. b_{nm} depends of course on $\psi_m^* \psi_n$.

Since it is certain that the atom will be in at least one of all the stationary states, the sum of all the probabilities $| c_n |^2$ must be a certainty. We get therefore:

$$\sum_n | c_n |^2 = 1. \tag{8}$$

10 · 2. *Calculation of Intensities, Selection and Polarisation Rules.*

Suppose that an electron is vibrating so that its displacement at any instant from its position of equilibrium is a vector $\vec{r}$ having the components x, y, z. Obviously, this represents an oscillating electric dipole of moment $\vec{p} = \epsilon \vec{r}$, having components $p_x = \epsilon x$, $p_y = \epsilon y$, $p_z = \epsilon z$.

Now according to the theory of Hertz, the field of such a dipole consists of an electric vector $\vec{E}$ and a magnetic vector $\vec{H}$, such that $\vec{E}$ is at right angles to $\vec{H}$, and both are perpendicular to the direction in which the radiation is propagated. Let θ be the angle between the direction of propagation of the radiation and the direction of r, and let ρ denote the distance from the centre of the dipole. Then Hertz's theory gives for the values of the vectors $\vec{E}$ and $\vec{H}$:

$$E = H = \frac{\sin\theta}{c^2\rho}\frac{d^2p}{dt^2}, \tag{1}$$

provided of course that ρ is very much greater than the wavelength of the radiation emitted.

From Poynting's theorem, we know that the energy radiated per second in a given direction is

$$\vec{S} = \frac{c}{4\pi}\left(\vec{E} \times \vec{H}\right),$$

where $\vec{E} \times \vec{H}$ denotes the vector product, whose value is given by EH sin (angle between $\vec{E}$ and $\vec{H}$). But since $\vec{E}$ is at right angles to $\vec{H}$, we get

$$S = \frac{c}{4\pi} EH. \tag{2}$$

From (1) and (2) we get therefore

$$\begin{aligned} S &= \frac{c}{4\pi}\cdot\left\{\frac{\sin\theta}{c^2\rho}\frac{d^2p}{dt^2}\right\}^2 = \frac{c}{4\pi}\left\{\frac{\sin\theta}{c^2\rho}\,\epsilon\,\frac{d^2r}{dt^2}\right\}^2 \\ &= \frac{\epsilon^2\sin^2\theta}{4\pi c^3\rho^2}\left(\frac{d^2r}{dt^2}\right)^2. \end{aligned} \tag{3}$$

The total energy I, radiated in unit time, is obtained by integrating S over the whole surface of a sphere σ:

$$\begin{aligned} I &= \iint S\,d\sigma \\ &= \frac{\epsilon^2}{4\pi c^3\rho^2}\left(\frac{d^2r}{dt^2}\right)^2 \int\limits_{\theta=0}^{\pi}\int\limits_{\phi=0}^{2\pi} \sin^2\theta\,\rho^2\sin\theta\,d\theta\,d\phi \\ &= \frac{\epsilon^2}{4\pi c^3}\left(\frac{d^2r}{dt^2}\right)^2\cdot 2\pi\cdot\int_0^{\pi}\sin^3\theta\,d\theta \\ &= \frac{2}{3}\frac{\epsilon^2}{c^3}\left(\frac{d^2r}{dt^2}\right)^2. \end{aligned} \tag{4}$$

The components of the intensity are given by :

$$I_x = \frac{2}{3}\frac{\epsilon^2}{c^3}\left(\frac{d^2x}{dt^2}\right)^2, \qquad (5)$$

with similar expressions for I_y and I_z.

With the help of the correspondence principle, this classical theory formula is taken over into the quantum theory, with the difference, however, that the radiation would now be due to a transition from the state *n* to the state *m*, so that *x, y, z* would now be matrices. For the intensity of radiation due to a transition from the state *n* to the state *m*, we would therefore have

$$\begin{aligned} I_x &= \frac{2\epsilon^2}{3c^3}\left(\frac{d^2x_{nm}}{dt^2}\right)^2 \\ I_y &= \frac{2\epsilon^2}{3c^3}\left(\frac{d^2y_{nm}}{dt^2}\right)^2 \qquad (6) \\ I_z &= \frac{2\epsilon^2}{3c^3}\left(\frac{d^2z_{nm}}{dt^2}\right)^2. \end{aligned}$$

But from (6) § 5 · 2, we know that time enters into the matrix elements according to

$$x_{nm} = x_{nm}\, e^{2\pi i\nu_{mn}t}$$

so that

$$\frac{d^2x_{nm}}{dt^2} = (2\pi i\nu_{nm})^2\, x_{nm}. \qquad (7)$$

Substituting (7) in the formulæ (6), we get for the intensity components :

$$\begin{aligned} I_x &= \frac{2}{3}\frac{\epsilon^2}{c^3}(2\pi i\nu_{nm})^4\, x_{nm}^2, \\ I_y &= \frac{2}{3}\frac{\epsilon^2}{c^3}(2\pi i\nu_{nm})^4\, y_{nm}^2, \qquad (8) \\ I_z &= \frac{2}{3}\frac{\epsilon^2}{c^3}(2\pi i\nu_{nm})^4\, z_{nm}^2. \end{aligned}$$

From (18) § 8 · 6 (2), we know that in the formulæ (8) above, the matrix elements x_{nm} are derived from the Schrödinger wave-functions according to the rule

$$x_{nm} = \int \psi_m^* x\, \psi_n\, dv, \qquad (9)$$

with similar expressions for y_{nm} and z_{nm}.

The expressions (8) enable us to find the selection and polarisation rules also. Thus, if it happens that for a certain transition $n \rightarrow m$, we get $x_{nm} = 0$, then we should expect no radiation of the frequency ν_{nm} polarised parallel to the *x*-axis.

To get the selection rules, therefore, we must find out under what conditions $\int \psi_m^* x \psi_n \, dv$ will vanish.

Moreover, the polarisation of the spectral line with the frequency ν_{nm} is determined from the values and relative phases of x_{nm}, y_{nm} and z_{nm}.

10 · 3. *Complimentary Character of the Corpuscular and the Wave Conceptions.*

In the foregoing section we saw how the two ideas of corpuscles and waves can be successfully merged into a statistical interpretation. There is thus no essential contradiction between the two conceptions, as employed in quantum mechanics, nor are they mutually exclusive. As pointed out by Niels Bohr in 1927, the one representation supplements the other. "Particle" as well as "wave" are each just half the reality underlying both matter and radiation. The word "wavicle" has been coined by Eddington to describe these ultimate constituents.

A very illuminating analysis of the whole problem has been given by Dirac with the help of the "Principle of Superposition". In this connection, Dirac also raises the question of the aim and object of theoretical physics, and supplies the answer that its only object is to calculate results that can be compared with experiments, and not to give more detailed answers than can be experimentally verified.

For explaining the principle of superposition, Dirac considers the simple case of the polarisation of light. Experiments on the photo-electric effect show that if the incident light is plane polarised, the photo-electrons are ejected in a preferential direction. This leads us to the conclusion that "the polarisation (= wave) properties of light are closely connected with its corpuscular properties, and that one must ascribe a polarisation to the photons. One must consider, for instance, a beam of light plane polarised in a certain direction as consisting of photons each of which is plane polarised in that direction, and a beam of circularly polarised light as consisting of photons each circularly polarised."

For these considerations to hold, it is necessary that a certain state of polarisation is ascribed to every photon. With the help of this idea of the state of polarisation of a photon, Dirac gives a description of the experimental facts about the resolution of light into polarised components, and the recomposition of these components.

He considers a beam of light passing through a crystal of tourmaline, which has the property of letting through only light which is plane polarised perpendicular to its optic axis. Now the classical theory predicts that if the beam is polarised perpendicular to the axis, it will all go through the crystal; if it is polarised parallel to the axis, none of it will go through; finally if it is polarised at an angle α to the axis, a fraction $\sin^2 \alpha$ will go through.

Accordingly, Dirac assumes that each photon polarised perpendicular to the axis passes through the crystal, while each photon polarised parallel to the axis is absorbed. In the case of the obliquely polarised incident beam, each of the incident photons is obliquely polarised at the same angle α. Now, if we perform the experiment of letting an individual photon pass through the crystal, then according to quantum mechanics, sometimes we shall find that the whole photon passes through the crystal unhindered, and at other times we shall find that the whole photon is absorbed. We shall never find that part of the photon passes through, and part of it is absorbed. If we perform the experiment a large number of times, we shall find that in a fraction $\sin^2 \alpha$ of the total number of times the photon passes through the crystal, and in a fraction $\cos^2 \alpha$ times the photon is absorbed. This fact is interpreted by Dirac to mean that the photon has a probability $\sin^2 \alpha$ of passing through the crystal, and a probability $\cos^2 \alpha$ of being absorbed. These values for the probabilities lead to the correct classical results for an incident beam containing a large number of photons. The individuality of the photon is thus preserved in all cases.

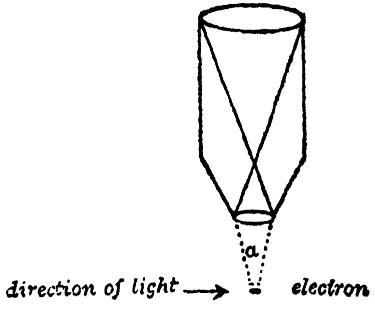
To provide a further description of this fact, Dirac agrees to regard an obliquely polarised photon, as being partly in the state of polarisation parallel to the axis, and partly in a state of polarisation perpendicular to the axis. He considers the state of oblique polarisation as a kind of superposition of the two states of parallel and perpendicular polarisation. Any state of polarisation is thus assumed to be capable of being resolved into any two mutually perpendicular states of polarisation.

Thus, before the photon is observed, it is partly in each of the two states of parallel and perpendicular polarisations. When we make an observation on it by passing it through the crystal, it is forced entirely into the state of parallel, or entirely into the state of perpendicular polarisation, and is thus observed either to pass through the crystal or to be absorbed by it. Which of the two states it will be forced into cannot be predicted exactly, but is governed only by probability laws.

Thus in trying to reconcile the corpuscular and wave conceptions, we are led necessarily to an uncertainty about the result of an experiment. At the most we can only predict the probability of occurrence of each of the possible results. This indeterminacy in modern physics was first elucidated in 1927 by W. Heisenberg, who considered the experiments that can be performed to determine the position and velocity of an electron, and proved that it is impossible to determine simultaneously both the position and the velocity of the electron with exactness. Thus Heisenberg formulated the principle that a certain indeterminacy is necessarily associated with all measurements of atomic magnitudes. It arises from the fact that every observation necessarily disturbs the thing observed, but the extent of this disturbance itself is quite undetermined. We shall illustrate this point by describing in detail two experiments, one to fix the position of an electron and another to determine its momentum.

10 · 4. *Determination of the Position of a Free Electron.*

In order to fix the position of the electron as accurately as possible, we must illuminate it with the light of as small wavelength as possible. For this purpose we shall assume with Heisenberg that we are employing a γ-ray microscope.[1]



---

[1] Prof. Heisenberg himself told the author that he had thought of the γ-ray microscope in 1924, *i.e.*, before he discovered the quantum mechanics, and long before he published the uncertainty principle.

Let the cone of rays scattered from the electron through the objective have an angular opening α, and let the wave-length of the light illuminating it be λ. Now it is known from the laws of optics that the microscope cannot resolve distances of less than the amount $\frac{\lambda}{\sin\alpha}$. This is called the resolving power of the microscope. Thus if x is the direction in which the light falls on the electron, the uncertainty in the knowledge of the position of the electron is of the order of $\frac{\lambda}{\sin\alpha}$, so that

$$\Delta x \sim \frac{\lambda}{\sin\alpha}. \tag{1}$$

From the Compton-effect, however, we know that when a photon collides with an electron and is scattered from it, the electron also receives a recoil, the value of the momentum depending on the direction of the scattered photon. The scattered photon can take any direction lying in the cone of rays of angle α, and consequently the recoil momentum may lie anywhere between $\frac{h}{\lambda}(1 - \frac{1}{2}\sin\alpha)$ and $\frac{h}{\lambda}(1 + \frac{1}{2}\sin\alpha)$. Thus, if p_x is the component of momentum in the direction of the x-axis, then there is an uncertainty in our knowledge of the momentum of the order

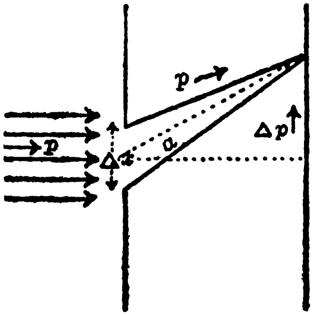
$$\Delta p_x \sim \frac{h}{\lambda}\sin\alpha. \tag{2}$$

The product of the two uncertainties is therefore

$$\Delta x\, \Delta p_x \sim h. \tag{3}$$

This is the famous "uncertainty relation" of Heisenberg. It asserts that the position and momentum cannot be exactly determined simultaneously; the product of their uncertainties must always be of the order of Planck's constant.

On account of the importance of the uncertainty relation, we shall prove its validity in yet another experiment, that of the diffraction of the electrons at a slit, which can be employed to measure the position of an electron. The following account is substantially that given by Heisenberg and Born.



Let a number of electrons pass through a slit of width Δx, and produce a diffraction pattern on a screen placed to the right, as in the accompanying figure. The fact that a diffraction pattern appears merely allows us to assert that each individual electron has passed through the slit; we cannot definitely say at what place in the slit the passage actually took place. Thus there is an uncertainty of the amount Δx in the specification of position perpendicular to the direction of incidence of the electrons.

After passing through the slit, the individual electron suffers deflection, and acquires a component momentum Δp, perpendicular to its original direction of incidence. From the figure it is apparent that the mean value of Δp is given by

$$\Delta p \sim p \sin \alpha, \tag{4}$$

where α is the mean angle of deflection.

The above account is based on the corpuscluar point of view. But the wave-theory of matter asserts that the electron may also be considered as a de Broglie-wave of the wave-length λ given by

$$\lambda = \frac{h}{p}, \quad \text{or } p = \frac{h}{\lambda}, \tag{5}$$

where p is the momentum of the electron. Then from the laws of optics, we know that the angle of deflection α is connected with the wave-length λ and the width Δx of the slit by the relation

$$\sin \alpha \sim \frac{\lambda}{\Delta x}. \tag{6}$$

Substituting (5) and (6) in (4), we get

$$\Delta p \sim p \sin \alpha \sim \frac{h}{\lambda} \cdot \frac{\lambda}{\Delta x} = \frac{h}{\Delta x},$$

so that, we have again the uncertainty relation:

$$\Delta p \, \Delta x \sim h. \tag{7}$$

There is thus a complete reciprocity between our knowledge of the position and that of the momentum or velocity of the electron. The more exactly we know the position, the less precise would be our knowledge of the velocity, and *vice versa*.

For the sake of convenience, we have taken only one co-ordinate x. Obviously, the argument applies to any other co-ordinate.

10 · 5. *Determination of the Velocity of an Electron.*

In this section we shall verify the uncertainty relation for an experiment performed to measure the velocity of an electron.

Suppose that light of frequency ν is falling on the electron in the direction of the x-axis, and that the scattered light is observed in the direction of the y-axis. We assume that the component p_x of the electron's momentum is known quite exactly, so that according to the uncertainty relation, nothing whatever is known about the x-co-ordinate. On the other hand, we assume that the y-co-ordinate of the electron is known quite accurately, so that we are completely ignorant about the component of momentum p_y.

When in the experiment we observe the scattered light in the direction of the y-axis, we gain some knowledge of the y-co-ordinate of the electron. The experiment, however, introduces an element of uncertainty in our knowledge of the component of momentum p_y to the extent that the product of the uncertainties will satisfy Heisenberg's relation.

We make use of the Doppler effect to prove this. Let E denote the energy of the electron and ν the frequency of light before the collision. Similarly, let E′ denote the energy of the electron and ν' the frequency of light after the collision. Remembering that the direction of the incident light is x and that of the scattered light y, and applying the laws of conservation of energy and momentum to the system (consisting of an electron and a photon), we get the equations

$$E + h\nu = E' + h\nu', \tag{1}$$

$$p_x + \frac{h\nu}{c} = p_x', \tag{2}$$

$$p_y = p_y' + \frac{h\nu'}{c}, \tag{3}$$

where p_x', p_y' are the components of momentum of the electron after the collision.

Let m be the mass of the electron, then we have

$$E = \frac{1}{2m}(p_x^2 + p_y^2), \quad E' = \frac{1}{2m}(p_x'^2 + p_y'^2). \tag{4}$$

Thus from (1) and (4), we get

$$\begin{aligned} h(\nu - \nu') &= E' - E \\ &= \frac{1}{2m}\left\{(p_x'^2 + p_y'^2) - (p_x^2 + p_y^2)\right\} \\ &= \frac{1}{2m}\left\{(p_x'^2 - p_x^2) + (p_y'^2 - p_y^2)\right\} \\ &= \frac{1}{2m}\left\{(p_x' - p_x)(p_x' + p_x) + (p_y' - p_y)(p_y' + p_y)\right\}. \end{aligned}$$

But the difference $p_x' - p_x$ is very small compared to p_x, and similarly $p_y' - p_y$ is very small compared to p_y, we can write

$$p_x' + p_x \sim 2p_x, \quad p_y' + p_y \sim 2p_y. \tag{5}$$

Thus

$$h(\nu - \nu') \sim \frac{1}{m}\left\{(p_x' - p_x)\,p_x + (p_y' - p_y)\,p_y\right\}.$$

Substituting in this from (2) and (3), we get

$$\begin{aligned} h(\nu - \nu') &\sim \frac{1}{m}\left\{\frac{h\nu}{c}p_x - \frac{h\nu'}{c}p_y\right\} \\ &\sim \frac{h\nu}{mc}(p_x - p_y), \end{aligned} \tag{6}$$

since ν' is of the same order of magnitude as ν. In the relation (6) all quantities except p_y and ν' are assumed to be known completely ; so that the accuracy of our knowledge of p_y depends only on the accuracy of the knowledge of ν', thus giving us, on taking the finite differences of both sides,

$$h\Delta\nu' \sim \frac{h\nu}{mc}\Delta p_y$$

or

$$\Delta p_y \sim \frac{mc}{\nu}\Delta\nu'. \tag{7}$$

To determine ν' with this accuracy, it is necessary to observe a train of waves of finite length, which in turn, demands a finite time T given by

$$T = \frac{1}{\Delta\nu'}, \tag{8}$$

We can see that this finite interval of time causes an uncertainty in our knowledge of the position of the electron. For we do not know whether the photon collided with the electron at the beginning or at the end of this time interval, so that we cannot say whether the velocity of the electron was $\frac{p_y}{m}$ or $\frac{p_y'}{m}$ during this time. The uncertainty in the position of the electron which is produced by this cause is thus

$$\Delta y = \frac{1}{m}(p_y - p_y')\,\mathrm{T}.$$

Substituting the value of $p_y - p_y'$ from (3), we get

$$\Delta y = \frac{h\nu'}{mc}\mathrm{T} \sim \frac{h\nu}{mc}\mathrm{T}$$

$$\sim \frac{h\nu}{mc}\cdot\frac{1}{\Delta\nu'}. \tag{9}$$

From (7) and (9), we get finally

$$\Delta y\,\Delta p_y \sim \frac{h\nu}{mc}\,\frac{1}{\Delta\nu'} \times \frac{mc}{\nu}\,\Delta\nu' \tag{10}$$

$$\sim h.$$

This proves the desired relation.

10·6. *The Uncertainty Relation. Theory of Operators.*

In the last two sections, we have shown by a consideration of special cases that the product of the uncertainties in the position and in the velocity is of the order of magnitude h.

We can, however, deduce the exact relation between these uncertainties without making any use of the corpuscular or wave pictures. It appears then that the uncertainty relation is a logical consequence of the quantum condition :

$$pq - qp = \frac{h}{2\pi i}\,1 \tag{1}$$

and of the mathematical apparatus of the quantum mechanics.

For the sake of convenience, we shall confine ourselves to the case of a single degree of freedom characterised by the co-ordinate q and the momentum p. The general case brings about no essential change in the argument.

From the theory of wave-mechanics, we have the Schrödinger wave-equation

$$(\mathrm{H} - \mathrm{E})\,\psi = 0 \text{ or } \mathrm{H}\psi = \mathrm{E}\psi, \tag{2}$$

where $\psi(q)$ is the wave-function, and $H(p, q)$ is the Hamiltonian operator, p standing for

$$p = \frac{h}{2\pi i}\frac{\partial}{\partial q}. \tag{3}$$

We have seen also that for $E < 0$, solutions of the equation (2) exist only for certain discrete values E_n, which are called the eigen-values, the corresponding solutions ψ_n being called the eigen-functions :

$$H\psi_n = E_n\psi_n \quad (n = 1, 2, 3, \cdots) \tag{4}$$

Generalising this idea, we take any function $A(p, q)$ as a linear operator, which acting upon a wave-function $\psi(q)$ transforms it into another wave-function $\phi(q)$, and express this fact by :

$$A\psi = \phi. \tag{5}$$

Special cases of (5) are provided by taking $A = q$ and $A = p$. $q\psi$ means simply multiplication by q, and $p\psi$ means, on account of (3), $\frac{h}{2\pi i}\frac{\partial\psi}{\partial q}$. Since the operators are not mere numbers, we can see that in general

$$AB \neq BA.$$

A particular example of this non-commutative property is given by the quantum condition from which we see that $pq \neq qp$.

Generally, the function ϕ will be quite different from ψ in (5). If, however, for any ψ_n it happens that the operator A reproduces it except for a constant factor a_n, *i.e.*, if

$$A\psi_n = a_n\psi_n, \tag{6}$$

then a_n is called an eigen-value and ψ_n an eigen-function of A.

The adjoint operator $A^\dagger$ is defined by the condition that if $\psi(q)$ and $\phi(q)$ are any two functions then we must have

$$\int\phi^*(A\psi)\,dq = \int(A^\dagger\phi)^*\cdot\psi dq, \tag{7}$$

where ϕ^* denotes the complex conjugate of the function ϕ, and where the integral extends over the whole domain of the co-ordinate q.

An operator A is called real if it is identical with its adjoint $A^\dagger$, *i.e.*, if

$$A = A^\dagger. \tag{8}$$

The co-ordinate q and the momentum p are easily seen to be real, according to this definition. For, we have

$$\int \phi^* (q\psi)\, dq = \int (q\phi^*)\, \psi dq = \int (q\phi)^*\, \psi dq,$$

so that $q = q^\dagger$. Similarly

$$\int \phi^* \cdot (p\psi)\, dq = \int \phi^* \frac{h}{2\pi i} \frac{\partial \psi}{\partial q}\, dq$$

$$= \frac{h}{2\pi i} [\phi^* \psi] - \int \frac{h}{2\pi i} \frac{\partial \phi^*}{\partial q} \cdot \psi\, dq.$$

Now the boundary condition would be always such that $[\phi^* \psi] = 0$, so that

$$\int \phi^* (p\psi) dq = - \int \left(\frac{h}{2\pi i} \frac{\partial \phi^*}{\partial q} \right) \psi\, dq$$

$$= \int (p\phi)^*\, \psi dq,$$

showing that $p = p^\dagger$, and that p is real.

It can be shown generally that the eigen-values of real operators are real For, in the equation (7) take $\phi = \psi = \psi_n$, then since $A^\dagger = A$, we get

$$\int \psi_n^* (A\psi_n)\, dq = \int (A\psi_n)^*\, \psi_n\, dq = \left\{ \int (A\psi_n) \cdot \psi_n^*\, dq \right\}^*, \tag{9}$$

i.e., the quantity $\int \psi_n^* (A\psi_n)\, dq$ is equal to its complex conjugate, so that it must be real. Now if a_n is the eigen-value belonging to the eigen-function ψ_n, then we have $A\psi_n = a_n \psi_n$. Hence

$$\int \psi_n^* (A\psi_n)\, dq = \int \psi_n^* (a_n \psi_n)\, dq = a_n \int \psi_n^*\, \psi_n\, dq,$$

so that

$$a_n = \frac{\int \psi_n^* (A\psi_n)\, dq}{\int |\psi_n|^2\, dq}. \tag{10}$$

In (9), we have proved that $\int \psi_n^* (A\psi_n)\, dq$ is real. Moreover, $|\psi_n|^2$ is also real, so that we see that a_n is also real.

Further, it is not difficult to prove that any two eigen-functions of a real operator A, belonging to different eigen-values are mutually orthogonal. For, if a_m, a_n are two eigen-values of a real operator A, and ψ_m, ψ_n the corresponding eigen-functions, we have

$$A\psi_m = a_m \psi_m, \quad A\psi_n = a_n \psi_n.$$

Taking the complex conjugate of the second equation:

$$(\mathrm{A}\,\psi_n)^* = a_n \psi_n^*.$$

Therefore we get on multiplying by ψ_n^* and ψ_m respectively:

$$\int \left\{ \psi_n^* (\mathrm{A}\,\psi_m) - \psi_m (\mathrm{A}\,\psi_n)^* \right\} dq = (a_m - a_n) \int \psi_m \psi_n^* \, dq \quad (11)$$

But since $\mathrm{A} = \mathrm{A}^\dagger$, we have from (7) on writing $\phi = \psi_n$, $\psi = \psi_m$:

$$\int \psi_n^* (\mathrm{A}\,\psi_m)\, dq = \int (\mathrm{A}^\dagger \psi_n)^* \psi_m \, dq = \int (\mathrm{A}\,\psi_n)^* \, \psi_m \cdot dq,$$

so that the left-hand side of (11) is zero and we get:

$$(a_m - a_n) \int \psi_m \psi_n^* \, dq = 0.$$

Since a_m and a_n are different, *i.e.*, since $a_m \neq a_n$, we get

$$\int \psi_m \psi_n^* \, dq = 0 \quad (m \neq n). \quad (12)$$

We assume that the eigen-functions are all normalised, *i.e.*,

$$\int \psi_n \psi_n^* \, dq = \int |\, \psi_n |^2 \, dq = 1. \quad (n = 1, 2, \cdots) \quad (13)$$

If this is not already the case, we can easily normalise them by multiplying ψ_n by a suitably chosen constant factor, which can be determined from the equation (13).

We can develop any function ϕ in a "Fourier series" of the eigen-functions ψ_n:

$$\phi = \sum_{n=1}^{\infty} c_n \psi_n, \; c_n = \int \phi \psi_n^* \, dq, \quad (14)$$

and its complex conjugate ϕ^* by the conjugate series:

$$\phi^* = \sum_{n=1}^{\infty} c_n^* \psi_n^*, \; c_n^* = \int \phi^* \psi_n \, dq. \quad (15)$$

Then we have

$$\int |\phi|^2 \, dq = \int \phi\phi^* \, dq = \int (\sum_n c_n \psi_n)\, \phi^* \, dq$$
$$= \sum_n c_n \int \psi_n \phi^* \, dq = \sum_n c_n c_n^* = \sum_n |\, c_n |^2. \quad (16)$$

Now we define the mean value $\bar{\mathrm{A}}$ of the quantity represented by the operator A in a state represented by the function ϕ by the equation:

$$\bar{\mathrm{A}} = \int \phi^* (\mathrm{A}\phi)\, dq = \int \phi^* (\sum_n c_n \, \mathrm{A}\,\psi_n)\, dq. \quad (17)$$

On account of $\mathrm{A}\,\psi_n = a_n \psi_n$ and (15) we get

$$\bar{\mathrm{A}} = \sum_n c_n \int \phi^* a \, \psi_n \, dq = \sum_n c_n a_n c_n^*$$
$$= \sum_n |\, c_n |^2 \, a_n.$$

On this definition it can be proved that the mean value of the product $AA^\dagger$ is never negative. For, writing $AA^\dagger$ in place of A in (17), we get

$$\overline{AA^\dagger} = \int \phi^* \cdot (AA^\dagger\phi)\, dq = \int \phi^* \cdot A\, (A^\dagger\phi)\, dq.$$

Let us write $A^\dagger\phi = \chi$, so that on account of (7) we have

$$\begin{aligned}\overline{AA^\dagger} &= \int \phi^* (A\chi)\ dq = \int (A^\dagger\phi)^*\, \chi\, dq \\ &= \int (A^\dagger\phi)^*\, (A^\dagger\phi)\, dq \\ &= \int |\, A^\dagger\phi\, |^2\, dq \geqslant 0. \qquad (19)\end{aligned}$$

We wish to determine the adjoint of the operator $(iA)^\dagger$. For this purpose, we multiply (7) by i, and get

$$i \int \phi^*\, (A\,\psi)\, dq = i \int (A^\dagger\phi)^*\ \psi\, dq$$

or since $(-i)^* = i$,

$$\int \phi^*\, (iA\psi)\, dq = i \int (-\ iA^\dagger\phi)^*\, \psi\, dq. \qquad (20)$$

On the other hand, considering (iA) itself as one operator, we get from the equation of definition (7):

$$\int \phi^*\, \{(iA)\,\psi\}\, dq = \int \{(iA)^\dagger\, \phi\}^*\ \psi\, dq. \qquad (21)$$

Comparing (20) with (21), we see that

$$(iA)^\dagger = -\, iA^\dagger. \qquad (22)$$

Obviously, therefore, we have for any two operators A and B:

$$(A + iB)^\dagger = A^\dagger - iB^\dagger. \qquad (23)$$

Suppose that A and B are real operators, and that λ is a real number. We have shown in (19) that the mean value of the product of any operator with its adjoint is never negative, so that we have:

$$\overline{(A+i\lambda B)\, (A + i\lambda B)^\dagger} \geqslant 0,$$

or, on account of (23) and because $A^\dagger = A$, $B^\dagger = B$:

$$\overline{(A + i\lambda B)\, (A - i\lambda B)} \geqslant 0,$$

or

$$f(\lambda) \equiv \overline{A^2} + \overline{B^2}\, \lambda^2 - i\lambda\, (\overline{AB - BA}) \geqslant 0. \qquad (24)$$

Since $f(\lambda)$ is real, we see that $\overline{AB - BA}$ is purely imaginary.

To determine the minimum value of $f(\lambda)$, we must take that value of λ which makes $\frac{df}{d\lambda}$ equal to zero, *i.e.*, the value for which

$$\frac{df}{d\lambda} = 2\,\overline{B^2}\, \lambda - i\, (\overline{AB - BA}) = 0,$$

Thus the minimum value of $f(\lambda)$ is given by

$$\lambda = \frac{i}{2}\,\frac{\overline{AB - BA}}{\overline{B^2}}. \tag{25}$$

Then

$$\{f(\lambda)\}_{\text{min.}} = \overline{A^2} + \overline{B^2}\left\{\frac{i}{2}\,\frac{\overline{AB - BA}}{\overline{B^2}}\right\}^2 - i\,(\overline{AB - BA})\cdot\left\{\frac{i}{2}\,\frac{\overline{AB - BA}}{\overline{B^2}}\right\}$$

$$= \overline{A^2} + \frac{1}{4}\,\frac{(\overline{AB - BA})^2}{\overline{B^2}} \geqslant 0. \tag{26}$$

Hence we get :

$$\overline{A^2}\;\overline{B^2} \geqslant -\frac{1}{4}\,(\overline{AB - BA})^2. \tag{27}$$

It must be remembered that $\overline{AB - BA}$ is purely imaginary, and therefore the right-hand side of (27) is really positive.

Now let

$$\delta A = A - \bar{A}, \quad \delta B = B - \bar{B}, \tag{28}$$

Then

$$\begin{aligned}\delta A\cdot\delta B - \delta B\cdot\delta A &= (A - \bar{A})\,(B - \bar{B}) - (B - \bar{B})\,(A - \bar{A})\\ &= AB - A\bar{B} - \bar{A}B + \bar{A}\bar{B}\\ &\qquad - BA + B\bar{A} + \bar{B}A - \bar{B}\,\bar{A}\\ &= AB - BA,\end{aligned} \tag{29}$$

because $\bar{A}$ and $\bar{B}$ are not operators but ordinary numbers, and therefore $\bar{A}\bar{B} = \bar{B}\bar{A}$, both being the same operator.

Since the relation (27) is true for any two operators, we replace A by δA and B by δB :

$$\overline{(\delta A)^2}\cdot\overline{(\delta B)^2} \geqslant -\frac{1}{4}\,\{\overline{(\delta A)\,(\delta B) - (\delta B)\,(\delta A)}\}^2,$$

or, on account of (29) :

$$\overline{(\delta A)^2}\cdot\overline{(\delta B)^2} \geqslant -\frac{1}{4}\,(\overline{AB - BA})^2. \tag{30}$$

For A and B we take the momentum p and the co-ordinate q respectively. Then we have :

$$\overline{(\delta p)^2}\cdot\overline{(\delta q)^2} \geqslant -\frac{1}{4}\,(\overline{pq - qp})^2. \tag{31}$$

But from the quantum condition (1), *viz.*, $pq - qp = \frac{h}{2\pi i}\,1$, we see that $(\overline{pq - qp})^2 = \left(\frac{h}{2\pi i}\right)^2 = -\frac{h^2}{4\pi^2}$, so that

$$\overline{(\delta p)^2}\cdot\overline{(\delta q)^2} \geqslant \frac{h^2}{16\pi^2}. \tag{32}$$

We define the uncertainty in any quantity to be the square root of the mean value of the square of its deviation, so that

$$\Delta p = \{\overline{(\delta p)^2}\}^{\frac{1}{2}}, \quad \Delta q = \{\overline{(\delta q)^2}\}^{\frac{1}{2}}, \tag{33}$$

then we get from (32) :

$$\Delta p \cdot \Delta q \geqslant \frac{h}{4\pi}. \tag{34}$$

This is then the exact determination of Heisenberg's uncertainty relation, and it is true for any two corresponding conjugate quantities.

Since the energy E and time t are also corresponding conjugate quantities, we would have accordingly

$$\Delta E \cdot \Delta t \geqslant \frac{h}{4\pi}. \tag{35}$$

10 · 7. *Principle of Indeterminacy and the Concept of Causality.*

The uncertainty relations force us to the conclusions that it is impossible simultaneously to carry out the measurement of both the position and the velocity with any desirable accuracy. All our observations, therefore, represent only a selection out of a much broader range of possible observations, so that we get only probabilities of occurrences. Consequently, it is impossible to predict with certainty the result of any experiment.

This conclusion is, however, in complete contradiction to the " Law of Causality, " as understood in Classical Physics. According to this law, the behaviour of a system is subject to a rigid determinism. If the initial conditions of a system are given, then we can calculate the state of the system at any subsequent time and predict it with any desired accuracy.

With his uncertainty relations, Heisenberg has shown that this law of causality in the classical sense, with its consequent determinism, is no more valid. It can be said that if we know the wave-function ψ at any time we can calculate exactly its value at a subsequent time from Schrödinger's equation. It would seem therefore that the wave-function ψ can be accurately determined from its initial value. And since every physical quantity depends on ψ, the system may behave, after all, in a strictly causal way.

A little reflection, however, convinces us that this conception of causality is completely empty, and is, therefore, quite useless for our purpose. The quantity, which has physical significance, is $|\psi|^2$, so that initially, we can know only the value of $|\psi|^2$. ψ, being a complex quantity, cannot be completely determined from the value of $|\psi|^2$. Thus, as Heisenberg has pointed out, it may be true that "if we know the present exactly, we can predict the future," but it is impossible to know the present exactly. Here is the break from classical determinism. The world processes do take place in a strictly causal way, but we cannot predict them, because we cannot know the initial conditions exactly. This indeterministic view prevails among the majority of the physicists, its chief exponents being Bohr, Heisenberg, Born, Eddington, Dirac and others. They are of the opinion that all physical laws are statistical laws, and that physics should be based on the calculus of probability.

There is, however, an opposite school of thought, led by equally eminent physicists like Planck and Einstein, which seeks to modify the concept of causality so that it may still remain valid in the new philosophy of Science. For this school, the question of the law of causality being right or wrong is meaningless. It should rather be judged on the merits of its utility. Planck considers the law to be an extremely valuable guiding principle of human knowledge. He tries therefore to modify the concept of causality so as to reconcile it with the uncertainty relations.

For this purpose, he remarks that the uncertainty of a measurement is due to the fact that the amount of the quantity to be measured depends on the nature of the measuring process. He finds it, therefore, necessary to include the observer and the apparatus in the system which is being measured. When this is done, all laws of nature become again strictly causal. The notion of probability arises only when no account is taken of the measuring apparatus.

There is, however, a serious objection to this view. It makes the relation of the causal connection between two successive events dependent on the human mind. This anthropomorphism would give quite a provisional and changeable character to such a fundamental physical concept as the law of causality.

To remove this defect, the determinist introduces "an ideal mind, apprehending everywhere all the physical occurrences in their minutest points". Such an ideal mind would be able to predict an event with absolute accuracy. This explains why it is actually impossible to predict an event accurately in an individual case. "It is a natural consequence of the circumstance that man with his senses, and his measuring instruments is a part of nature. He is subject to her laws and cannot escape from her, while such a tie does not exist for the ideal mind." (Planck.)

Obviously this leads the discussion into the realm of metaphysics, and does not concern us here. The undisputed fact remains that in scientific investigations and physical description of nature there is no escape for us from Heisenberg's uncertainty relations.

CHAPTER XI.

RELATIVISTIC QUANTUM MECHANICS.

11·1 (1). *Conservation of Electric Charge.*

ACCORDING to modern ideas, an electric charge is essentially a quantity having discrete magnitude, which is determined by counting the number of electrons and protons. The total number of electrons and protons cannot naturally depend on the motion of the observers. We must therefore have the electric charge invariant for all observers.

Starting from this assumption of the invariance of electric charge, we shall demonstrate that the electro-magnetic field equations satisfy the principle of relativity. Thus we should have no reason to modify these equations.

11·1 (2). *Density of Electric Charge and Electric Current.*

First we shall find the laws of transformation for density and current. For definiteness we assume all electric charges to be composed of elementary charges, each of amount e.

In the neighbourhood of a specified point P of space, let $n(u)$ be the number of these charges per unit volume moving with velocity $\vec{u}$, as observed by S. Let n_0 be the number of the same charges per unit volume, as observed by S_0 who is moving with them. Then S sees this latter volume as contracted by a factor $\sqrt{1-\frac{u^2}{c^2}}$ in the direction of its motion, and therefore estimates the density as $\frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$ times greater than does S_0, so that we find

$$n(u) = \frac{n_0}{\sqrt{1-\frac{u^2}{c^2}}}. \tag{1}$$

In the frame K of the observer S, the charges under consideration will contribute an amount $en(u)$ to the charge-density ρ. The number of elementary charges which cross a unit area placed perpendicular to the x-axis per unit time is $u_x n(u)$, so that they contribute an amount $e\, u_x\, n\,(u)$ to the x-component of the current-density $\vec{\sigma}$. Similarly for the other components. Hence in K, we have

$$\rho = \sum_u en\,(u), \quad \vec{\sigma} = \sum_u e\, \vec{u}\, n\,(u), \tag{2}$$

where the summation is taken over all the velocities u which occur.

Now consider an observer S′ at rest in a frame K′ which is moving relative to K with a velocity v in the direction of the positive x-axis. Then S′ obtains similarly for the charges considered in (1) a number-density $n'\,(u')$ given by

$$n'\,(u') = \frac{n_0}{\sqrt{1 - \dfrac{u'^2}{c^2}}}, \tag{3}$$

where u' is the velocity of the charges as observed by S′.

From the equations (3) and (5) of § 1·6 (1), we get the identity :

$$\frac{1}{\sqrt{1 - \dfrac{u'^2}{c^2}}} = \frac{\gamma\,(1 - u_x v/c^2)}{\sqrt{1 - \dfrac{u^2}{c^2}}}, \quad \gamma = \frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \tag{4}$$

Therefore on account of (1) and (3) we obtain

$$\begin{aligned} n'\,(u') &= \gamma\,(1 - u_x v/c^2)\frac{n_0}{\sqrt{1 - \dfrac{u^2}{c^2}}} \\ &= \gamma\left(1 - \frac{u_x v}{c^2}\right) n\,(u). \end{aligned} \tag{5}$$

But from the Lorentz-transformation

$$dx' = \gamma\,(dx - vdt),\ dy' = dy,\ dz' = dz,\ dt' = \gamma\left(dt - \frac{v}{c^2}\,dx\right),$$

we find the law of transformation of velocity :

$$\begin{aligned} u_x' &= \frac{dx'}{dt'} = \frac{\gamma\,(dx - v\,dt)}{\gamma\left(dt - \dfrac{v}{c^2}\,dx\right)} = \frac{u_x - v}{1 - \dfrac{u_x v}{c^2}}, \quad \left[u_x = \frac{dx}{dt}\right], \\ u_y' &= \frac{dy'}{dt'} = \frac{dy}{\gamma\left(dt - \dfrac{v}{c^2}\,dx\right)} = \frac{u_y}{\gamma\left(1 - \dfrac{u_x v}{c^2}\right)}, \quad \left[u_y = \frac{dy}{dt}\right], \\ u_z' &= \frac{dz'}{dt'} = \frac{dz}{\gamma\left(dt - \dfrac{v}{c^2}\,dx\right)} = \frac{u_z}{\gamma\left(1 - \dfrac{u_x v}{c^2}\right)}, \quad \left[u_z = \frac{dz}{dt}\right]. \end{aligned} \tag{6}$$

Substituting these in (5), we obtain:

$$u_x' n'(u') = \frac{u_x - v}{1 - \frac{u_x v}{c^2}} \cdot \gamma \left(1 - \frac{u_x v}{c^2}\right) n(u) = \gamma (u_x - v)\, n(u).$$

$$u_y' n'(u') = u_y n(u), \; u_z' n'(u) = u_z n(u). \tag{7}$$

Hence, the charge and current densities in the frame K′ are given by

$$\begin{aligned} \rho' &= \Sigma\, e n'(u') = \Sigma\, e \gamma \left(1 - \frac{u_x v}{c^2}\right) n(u) \\ &= \gamma \left(\rho - \sigma_x \frac{v}{c^2}\right); \end{aligned} \tag{8}$$

$$\begin{aligned} \sigma_x' &= \Sigma\, e u_x' n'(u') = \Sigma\, e\gamma (u_x - v)\, n(u) = \gamma (\sigma_x - v\rho), \\ \sigma_y' &= \Sigma\, e u_y' n'(u') = \Sigma\, e u_y n(u) = \sigma_y, \\ \sigma_z' &= \Sigma\, e u_z' n'(u') = \Sigma\, e u_z n(u) = \sigma_z. \end{aligned} \tag{9}$$

Thus the law of transformation for $(\sigma_x, \sigma_y, \sigma_z, \rho)$ is the same as the Lorentz-transformation for (x, y, z, t). Consequently, the inverse transformation would be

$$\sigma_x = \gamma (\sigma_x' + v\rho), \; \sigma_y = \sigma_y', \; \sigma_z = \sigma_z', \; \rho = \gamma \left(\rho' + \sigma_x' \frac{v}{c^2}\right). \tag{10}$$

11·1 (3). *The Maxwell-Lorentz Field Equations.*

From classical electrodynamics, we know that the Maxwell-Lorentz field equations are

$$\operatorname{div} \vec{E} = \rho, \quad \operatorname{div} \vec{H} = 0, \tag{1}$$

$$\operatorname{curl} \vec{E} = -\frac{1}{c} \frac{\partial \vec{H}}{\partial t}, \quad \operatorname{curl} \vec{H} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t} + \rho \frac{\vec{u}}{c}, \tag{2}$$

where $\vec{E}$ and $\vec{H}$ are the intensities of the electric and magnetic fields, ρ the density of the electric charge, and $\vec{u}$ the velocity with which it is moving.

From the equation (2) we obtain, since $\vec{\sigma} = \rho \vec{u}$,

$$\operatorname{div} \operatorname{curl} \vec{H} = \frac{1}{c} \frac{\partial}{\partial t} (\operatorname{div} \vec{E}) + \frac{1}{c} \operatorname{div} \vec{\sigma}.$$

But the divergence of a curl is identically zero, therefore substituting $\operatorname{div} \vec{E} = \rho$ from (1) we obtain:

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{\sigma} \equiv \frac{\partial \rho}{\partial t} + \frac{\partial \sigma_x}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \sigma_z}{\partial z} = 0. \tag{3}$$

This is the equation of continuity which expresses the conservation of total electric charge.

Further, if we consider the field equations for the case of free space with $\rho = 0$, then from (2) we get:

$$\text{curl curl } \vec{E} = -\frac{1}{c}\frac{\partial}{\partial t}\text{ curl } \vec{H} = -\frac{1}{c}\frac{\partial}{\partial t}\left\{\frac{1}{c}\frac{\partial \vec{E}}{\partial t}\right\}.$$

But we know from vector analysis:

$$\text{curl curl } \vec{E} = \text{grad div } \vec{E} - \nabla^2 \vec{E}, \tag{4}$$

where ∇^2 is the Laplacian operator $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$.

From the last two equations we obtain, since div $\vec{E} = \rho = 0$,

$$-\nabla^2 \vec{E} = -\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2}, \text{ or } \Box\vec{E} = 0, \tag{5}$$

where $\Box$ is the D'Alumbertian $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$.

Similarly, for the magnetic field $\vec{H}$, we find

$$\Box\vec{H} \equiv \nabla^2 \vec{H} - \frac{1}{c^2}\frac{\partial^2 \vec{H}}{\partial t^2} = 0. \tag{6}$$

These are the well-known equations for the propagation of Hertzian waves in free space, with the velocity c.

In the general case, we know from electrodynamics that the fields $\vec{E}$ and $\vec{H}$ can be determined in terms of the scalar potential ϕ and the vector potential $\vec{A}$, by means of the equations

$$\vec{E} = -\text{grad}\,\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}, \tag{7}$$

$$\vec{H} = \text{curl } \vec{A}. \tag{8}$$

But these equations do not determine the electromagnetic potentials A_x, A_y, A_z, ϕ uniquely. The potentials are concerned in the actual phenomena only through their curl, *viz.*, the electromagnetic force. The curl is unaltered if we replace $-A_x$ by $-A_x + \frac{\partial V}{\partial x}$, $-A_y$ by $-A_y + \frac{\partial V}{\partial y}$, $-A_z$ by $-A_z + \frac{\partial V}{\partial z}$, ϕ by $\phi + \frac{\partial V}{\partial t}$, where V is an arbitrary function of x, y, z, t. To avoid arbitrariness, an additional condition is imposed on the electromagnetic potentials :

$$\operatorname{div}\vec{A} + \frac{1}{c}\frac{\partial\phi}{\partial t} \equiv \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} + \frac{1}{c}\frac{\partial\phi}{\partial t} = 0. \tag{9}$$

Taking the divergence of both sides in (7) we obtain :

$$\operatorname{div}\vec{E} = -\operatorname{div}\operatorname{grad}\phi - \frac{1}{c}\frac{\partial}{\partial t}(\operatorname{div}\vec{A}).$$

But we have div grad $\phi = \nabla^2\phi$, div $\vec{E} = \rho$; div $\vec{A} = -\frac{1}{c}\frac{\partial\phi}{\partial t}$,

Therefore

$$\nabla^2\phi - \frac{1}{\partial^2}\frac{\partial^2\phi}{\partial t^2} = -\rho \text{ or } \Box\,\phi = -\rho. \tag{10}$$

From equation (8) we obtain :

$$\operatorname{curl}\vec{H} = \operatorname{curl}\operatorname{curl}\vec{A} = \operatorname{grad}\operatorname{div}\vec{A} - \nabla^2\vec{A},$$

or

$$\frac{1}{c}\frac{\partial\vec{E}}{\partial t} + \frac{\vec{\sigma}}{c} = \operatorname{grad}\left(-\frac{1}{c}\frac{\partial\phi}{\partial t}\right) - \nabla^2\vec{A}$$

$$= \frac{1}{c}\frac{\partial}{\partial t}\left(\vec{E} + \frac{1}{c}\frac{\partial\vec{A}}{\partial t}\right) - \nabla^2\vec{A}.$$

Therefore

$$\nabla^2\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = -\frac{\vec{\sigma}}{c} \text{ or } \Box\,\vec{A} = -\frac{1}{c}\vec{\sigma}. \tag{11}$$

We shall now transform the equations (10) and (11) to the frame K′. In § 1·3 we have shown that the operator $\Box$ is an invariant, so that we have $\Box' = \Box$. From the transformation equations (10) § 11·1 (2), we get therefore

$$\Box' A_x = -\frac{\gamma}{c}(\sigma_x' + v\rho'),\ \Box' A_y = -\frac{1}{c}\sigma_y',\ \Box' A_z = -\frac{1}{c}\sigma_z',$$

$$\Box'\phi = -\gamma\left(\rho' + \frac{\sigma_x' v}{c^2}\right). \tag{12}$$

Multiplying $\Box'\ \phi$ by $\gamma\frac{v}{c}$ and $\Box'\ A_x$ by γ, and subtracting the first from the second, we obtain

$$\Box'\ \gamma\left(A_x - \frac{v\phi}{c}\right) = -\frac{1}{c}\sigma_x'. \tag{13}$$

Similarly, multiplying $\Box' \mathrm{A}_x$ by $\gamma \frac{v}{c}$ and $\Box \phi$ by γ, and subtracting the first from the second, we obtain

$$\Box' \gamma \left(\phi - \frac{v \mathrm{A}_x}{c}\right) = -\rho'. \tag{14}$$

Hence the equation (12) may be written in the form

$$\Box' \vec{\mathrm{A}}' = -\frac{1}{c} \vec{\sigma}', \quad \Box' \phi' = -\rho', \tag{15}$$

provided we set

$$\mathrm{A}'_x = \gamma\left(\mathrm{A}_x - \frac{v\phi}{c}\right), \quad \mathrm{A}_y' = \mathrm{A}_y, \quad \mathrm{A}_z' = \mathrm{A}_z, \quad \phi' = \gamma\left(\phi - \frac{v}{c}\mathrm{A}_x\right). \tag{16}$$

This again gives the law of transformation for $\left(\mathrm{A}_x, \mathrm{A}_y, \mathrm{A}_z, \frac{\phi}{c}\right)$ of the same form as that for (x, y, z, t). The inverse transformation will be therefore

$$\mathrm{A}_x = \gamma\left(\mathrm{A}_x + \frac{v\phi'}{c}\right), \quad \mathrm{A}_y = \mathrm{A}_y', \quad \mathrm{A}_z = \mathrm{A}_z', \quad \phi = \gamma\left(\phi' + \frac{v}{c}\mathrm{A}_x'\right). \tag{17}$$

Substituting these values, we find that the equation (9) is transformed to

$$\operatorname{div}' \vec{\mathrm{A}}' + \frac{1}{c}\frac{\partial \phi'}{\partial t'} \equiv \frac{\partial \mathrm{A}_x'}{\partial x'} + \frac{\partial \mathrm{A}_y'}{\partial y'} + \frac{\partial \mathrm{A}_z'}{\partial z'} + \frac{1}{c}\frac{\partial \phi'}{\partial t'} = 0. \tag{18}$$

From (15) and (18) we see that the field equations hold unchanged in form in the system K′, when the potentials are transformed according to the law (16).

We now form the quantities $\vec{\mathrm{E}}'$, $\vec{\mathrm{H}}'$, given by

$$\vec{\mathrm{E}}' = -\operatorname{grad}' \phi' - \frac{1}{c}\frac{\partial \vec{\mathrm{A}}'}{\partial t'}, \quad \vec{\mathrm{H}}' = \operatorname{curl}' \vec{\mathrm{A}}'. \tag{19}$$

These $\vec{\mathrm{E}}'$ and $\vec{\mathrm{H}}'$ are the quantities recognised by S′ as the electric and magnetic intensities, because in his variables (x', y', z', t') they satisfy a set of equations in every respect of the same form as those satisfied by $\vec{\mathrm{E}}$ and $\vec{\mathrm{H}}$ for S.

The components of $\vec{\mathrm{E}}'$ and $\vec{\mathrm{H}}'$ can be evaluated by using (7), (8) and (16), together with the transformation equations for $\frac{\partial}{\partial x'}, \frac{\partial}{\partial y'}, \frac{\partial}{\partial z'}, \frac{\partial}{\partial t'}$. The latter are seen to be

$$\frac{\partial}{\partial x'} = \frac{\partial}{\partial x}\frac{\partial x}{\partial x'} + \frac{\partial}{\partial t}\frac{\partial t}{\partial x'} = \gamma\left(\frac{\partial}{\partial x} + \frac{v}{c^2}\frac{\partial}{\partial t}\right), \quad \frac{\partial}{\partial y'} = \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial z'} = \frac{\partial}{\partial z},$$

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial x}\frac{\partial x}{\partial t'} + \frac{\partial}{\partial t}\frac{\partial t}{\partial t'} = \gamma\left(\frac{\partial}{\partial t} + v\frac{\partial}{\partial x}\right).$$

Then from (19) we obtain:

$$E_x' = -\frac{\partial\phi'}{\partial z'} - \frac{1}{c}\frac{\partial A_x'}{\partial t'} = -\gamma\left(\frac{\partial}{\partial x} + \frac{v}{c^2}\frac{\partial}{\partial t}\right)\phi' - \frac{1}{c}\gamma\left(\frac{\partial}{\partial t} + v\frac{\partial}{\partial z}\right)A_x$$

$$= -\gamma\left(\frac{\partial}{\partial x} + \frac{v}{c^2}\frac{\partial}{\partial t}\right)\cdot\gamma\left(\phi - \frac{v}{c}A_x\right)$$

$$- \frac{1}{c}\gamma\left(\frac{\partial}{\partial t} + v\frac{\partial}{\partial x}\right)\gamma\left(A_x - \frac{v}{c}\phi\right)$$

$$= -\frac{\partial\phi}{\partial x} - \frac{1}{c}\frac{\partial A_x}{\partial t} = E_x.$$

$$H_z' = \frac{\partial A_y'}{\partial x'} - \frac{\partial A_x'}{\partial y'} = \gamma\left(\frac{\partial}{\partial x} + \frac{v}{c^2}\frac{\partial}{\partial t}\right)A_y - \frac{\partial}{\partial y}\cdot\gamma\left(A_x - \frac{v}{c}\phi\right)$$

$$= \gamma\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) + \gamma\frac{v}{c}\left(\frac{\partial\phi}{\partial y} + \frac{1}{c}\frac{\partial A_y}{\partial t}\right)$$

$$= \gamma\left(H_z - \frac{v}{c}E_y\right).$$

In this way, we find the law of transformation of $\vec{E}$ and $\vec{H}$ as follows:

$$\left.\begin{aligned} E_x' &= E_x,\ E_y' = \gamma\left(E_y - \frac{v}{c}H_z\right),\ E_z' = \gamma\left(E_z + \frac{v}{c}H_y\right);\\ H_x' &= H_x,\ H_y' = \gamma\left(H_y + \frac{v}{c}E_z\right),\ H_z' = \gamma\left(H_z - \frac{v}{c}E_y\right).\end{aligned}\right\} \qquad (20)$$

The inverse transformation is given by

$$\left.\begin{aligned} E_x &= E_x',\ E_y = \gamma\left(E_y' + \frac{v}{c}H_z'\right),\ E_z = \gamma\left(E_z' - \frac{v}{c}H_y'\right);\\ H_x &= H_x',\ H_y = \gamma\left(H_y' - \frac{v}{c}E_z'\right),\ H_z = \gamma\left(H_z' + \frac{v}{c}E_y'\right).\end{aligned}\right\} \qquad (21)$$

11·1 (4). *Relativistic Dynamics of a Charged Particle in an Electromagnetic Field.*

Suppose that the frame K is so chosen that the x-axis is parallel to the instantaneous velocity v of the charge e. Then in the frame K′ the charge is instantaneously at rest. We shall therefore assume that the usual electrostatic laws hold at that instant in K′ as far as the particle itself is concerned, so that the mechanical force acting upon it is

$$\vec{F}' = e\vec{E}'. \qquad (1)$$

In this case it is found that the components of the force $\vec{F}'$ are transformed according to the law:

$$F_x = F_x', \ F_y = \frac{1}{\gamma} F_y', \ F_z = \frac{1}{\gamma} F_z'.$$

Then using (20) § 11·1 (3), we find from (1):

$$\begin{aligned} F_x &= F_x' = eE_x' = eE_x, \\ F_y &= \frac{1}{\gamma} F_y' = \frac{1}{\gamma} eE_y' = \frac{1}{\gamma} e\cdot\gamma\left(E_y - \frac{v}{c} H_z\right), \\ F_z &= \frac{1}{\gamma} F_z' = \frac{1}{\gamma} eE_z' = \frac{1}{\gamma} e\cdot\gamma\left(E_z + \frac{v}{c} H_y\right). \end{aligned} \tag{2}$$

Since the velocity $\vec{u}$ of the particle has components $(v, 0, 0)$ as observed by S at the instant, the equations (2) can be written in the vector form

$$\vec{F} = e\left\{\vec{E} + \frac{1}{c}(\vec{u} \times \vec{H})\right\}, \tag{3}$$

where $\vec{u} \times \vec{H}$ denotes the vector product.

The equations of motion are given by $\frac{d\vec{p}}{dt} = \vec{F}$, so that

$$\begin{aligned} \frac{d}{dt}\left(\frac{m_0\vec{u}}{\sqrt{1-\frac{u^2}{c^2}}}\right) &= e\left\{\vec{E} + \frac{1}{c}(\vec{u} \times \vec{H})\right\} \\ &= e\left\{\left(-\operatorname{grad}\phi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t}\right) + \frac{1}{c}\left(\vec{u} \times \operatorname{curl}\vec{A}\right)\right\}, \end{aligned} \tag{4}$$

where m_0 is the rest-mass of the particle. The function T* is defined as in (6) § 2·4 by

$$T^* = m_0 c^2 \{(1 - \sqrt{1-\beta^2}\}, \ \beta = \frac{u}{c}. \tag{5}$$

Then the x-component of the equation of motion (4) is

$$\frac{d}{dt}\left(\frac{\partial T^*}{\partial \dot{x}}\right) = -e\frac{\partial\phi}{\partial x} - \frac{e}{c}\frac{\partial A_x}{\partial t} + \frac{e}{c}(\vec{u} \times \operatorname{curl}\vec{A})_x. \tag{6}$$

But

$$(\vec{u} \times \operatorname{curl}\vec{A})_x$$

$$= \dot{y}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - \dot{z}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)$$

$$= \left(\dot{x}\,\frac{\partial A_x}{\partial x} + \dot{y}\,\frac{\partial A_y}{\partial y} + \dot{z}\,\frac{\partial A_z}{\partial x}\right) - \left(\dot{x}\,\frac{\partial A_x}{\partial x} + y\,\frac{\partial A_x}{\partial y} + \dot{z}\,\frac{\partial A_x}{\partial z}\right)$$

$$= \frac{\partial}{\partial x}\left(\dot{x}\,A_x + \dot{y}\,A_y + \dot{z}\,A_z\right) - \frac{dA_x}{dt} + \frac{\partial A_x}{\partial t}.$$

Hence (6) may be written as

$$\frac{d}{dt}\left(\frac{\partial T^*}{\partial x} + \frac{e}{c}\,A_x\right) - \frac{\partial}{\partial x}\left\{- e\phi + \frac{e}{c}\,(\vec{u}\cdot\vec{A})\right\} = 0. \tag{7}$$

But if we consider the co-ordinates of position and velocities, *viz.*, $x, y, z, \dot{x}, \dot{y}, \dot{z}$, then T* does not depend on x, y, z explicitly, and ϕ does not depend on $x, \dot{y}, \dot{z}$, so that

$$\frac{\partial\phi}{\partial\dot{x}} = \frac{\partial\phi}{\partial\dot{y}} = \frac{\partial\phi}{\partial\dot{z}} = \frac{\partial T^*}{\partial x} = \frac{\partial T^*}{\partial y} = \frac{\partial T^*}{\partial z} = 0.$$

Also

$$A_x = \frac{\partial}{\partial\dot{x}}\left(\dot{x}\,A_x + \dot{y}\,A_y + \dot{z}\,A_z\right) = \frac{\partial}{\partial\dot{x}}\,(\vec{u}\cdot\vec{A}).$$

The equation of motion (7) can therefore be written as

$$\frac{d}{dt}\left\{\frac{\partial}{\partial\dot{x}}\left(T^* - e\phi + e\,\frac{\vec{u}\cdot\vec{A}}{c}\right)\right\} - \frac{\partial}{\partial x}\left\{T^* - e\phi + e\,\frac{\vec{u}\cdot\vec{A}}{c}\right\} = 0. \tag{8}$$

There are two similar equations for y and z. These are of the Lagrangian form with the Lagrangian function given by

$$L = T^* - e\phi + \frac{e}{c}\,(\vec{u}\cdot\vec{A}). \tag{9}$$

The momenta p_x, p_y, p_z are then

$$p_x = \frac{\partial L}{\partial\dot{x}} = \frac{\partial T^*}{\partial\dot{x}} + \frac{e}{c}\,A_x, \quad p_y = \frac{\partial T^*}{\partial\dot{y}} + \frac{e}{c}\,A_y, \quad p_z = \frac{\partial T^*}{\partial\dot{z}} + \frac{e}{c}\,A_z. \tag{10}$$

The Hamiltonian function H is then found to be

$$\begin{aligned} H &= \dot{x}\,p_x + \dot{y}p_y + \dot{z}\,p_z - L \\ &= \dot{x}\left(\frac{\partial T^*}{\partial\dot{x}} + \frac{e}{c}\,A_x\right) + \dot{y}\left(\frac{\partial T^*}{\partial\dot{y}} + \frac{e}{c}\,A_y\right) + \dot{z}\left(\frac{\partial T^*}{\partial\dot{z}} + \frac{e}{c}\,A_z\right) \\ &\quad - \left\{T^* - e\phi + \frac{e}{c}\,(\dot{x}\,A_x + \dot{y}\,A_y + \dot{z}\,A_z)\right\} \\ &= \dot{x}\,\frac{\partial T^*}{\partial\dot{x}} + \dot{y}\,\frac{\partial T^*}{\partial\dot{y}} + \dot{z}\,\frac{\partial T^*}{\partial\dot{z}} - T^* + e\phi. \end{aligned}$$

It was shown in (10) § 2·4 that

$$\sum_{xyz}\,\dot{x}\,\frac{\partial T^*}{\partial\dot{x}} - T^* = m_0c^2\left\{\frac{1}{\sqrt{1-\beta^2}} - 1\right\},$$

therefore we have

$$\mathrm{H} = m_0c^2\left\{\frac{1}{\sqrt{1-\beta^2}} - 1\right\} + e\phi. \tag{11}$$

To obtain the canonical form we require the expression of H in terms of p_x, p_y, p_z. For this, we see that

$$p_x - \frac{e}{c}\mathrm{A}_x = \frac{\partial \mathrm{T}^*}{\partial \dot{x}} = \frac{m_0\dot{x}}{\sqrt{1-\beta^2}}, \quad p_y - \frac{e}{c}\mathrm{A}_y = \frac{m_0\dot{y}}{\sqrt{1-\beta^2}},$$

$$p_z - \frac{e}{c}\mathrm{A}_z = \frac{m_0\dot{z}}{\sqrt{1-\beta^2}}.$$

Therefore

$$\Sigma\left(p_x - \frac{e}{c}\mathrm{A}_x\right)^2 = \frac{m_0{}^2u^2}{1-\dfrac{u^2}{c^2}},$$

so that

$$\Sigma\left(p_x - \frac{e}{c}\mathrm{A}_x\right)^2 + m_0{}^2c^2 = m_0{}^2\left(\frac{u^2}{1-\dfrac{u^2}{c^2}} + c^2\right) = \frac{m_0{}^2c^2}{1-\beta^2}.$$

Thus

$$\frac{m_0c^2}{\sqrt{1-\beta^2}} = c\left\{\Sigma\left(p_x - \frac{e}{c}\mathrm{A}_x\right)^2 + m_0{}^2c^2\right\}^{\frac{1}{2}}. \tag{12}$$

From (11) and (12) we obtain therefore

$$\mathrm{H} + m_0c^2 - e\phi = c\left\{\Sigma_{xyz}\left(p_x - \frac{e}{c}\mathrm{A}_x\right)^2 + m_0{}^2c^2\right\}^{\frac{1}{2}}. \tag{13}$$

If the particle were an electron, we would have to replace e by $-e$ throughout this section. We write also $\mathrm{E} = \mathrm{H} + m_0c^2$, E being the total energy, then from (13) we get for an electron

$$\mathrm{E} + e\phi = c\left\{\Sigma\left(p_x + \frac{e}{c}\mathrm{A}_x\right)^2 + m_0{}^2c^2\right\}^{\frac{1}{2}}, \tag{14}$$

or squaring,

$$\left(\frac{\mathrm{E}}{c} + \frac{e}{c}\phi\right)^2 = \Sigma_{xyz}\left(p_x + \frac{e}{c}\mathrm{A}_x\right)^2 + m_0{}^2c^2. \tag{15}$$

If there is no external field, we write $\phi = \mathrm{A}_x = \mathrm{A}_y = \mathrm{A}_z = 0$ in (13), and get the same Hamiltonian as (11) § 2·4.

If the velocity u of the electron can be taken to be very small compared to c, as assumed in the classical theory, then the term $m_0{}^2c^2$ under the root in (13) is very much larger than the rest.

Therefore, for the classical Hamiltonian H′ we get

$$\mathrm{H}' + m_0c^2 - e\phi = m_0c^2\left\{1 + \frac{1}{m_0{}^2c^2}\Sigma\left(p_x - \frac{e}{c}\mathrm{A}_x\right)^2\right\}^{\frac{1}{2}}$$

$$= m_0c^2\left\{1 + \frac{1}{2m_0{}^2c^2}\Sigma\left(p_x - \frac{e}{c}\mathrm{A}_x\right)^2\right\},$$

neglecting second and higher powers. Thus

$$\mathrm{H}' = e\phi + \frac{1}{2m_0}\sum_{xyz}\left(p_x - \frac{e}{c}\mathrm{A}_x\right)^2. \tag{16}$$

For the electron, we have therefore

$$\mathrm{H}' = -e\phi + \frac{1}{2m_0}\sum_{xyz}\left(p_x + \frac{e}{c}\mathrm{A}_x\right)^2. \tag{17}$$

11·2. *Earlier Attempts at a Relativistic Quantum Theory.*

The quantum mechanics developed so far has taken no account of the theory of relativity. This deficiency was supplied in the middle of the year 1926 by Schrödinger himself, and later by many other workers. Their contribution consisted in generalising the wave-equation to a relativistic form.

Thus, for a particle of mass m, Cartesian co-ordinates x, y, z and corresponding momenta p_x, p_y, p_z, moving under a conservative field of force of potential V, we had the classical Hamiltonian

$$\mathrm{H} = \frac{1}{2m}(p_x{}^2 + p_y{}^2 + p_z{}^2) + \mathrm{V}. \tag{1}$$

From this Hamiltonian, we derived the wave-equation by taking the momenta p_x, p_y, p_z, as operators defined by

$$p_x = \frac{h}{2\pi i}\frac{\partial}{\partial x}, \quad p_y = \frac{h}{2\pi i}\frac{\partial}{\partial y}, \quad p_z = \frac{h}{2\pi i}\frac{\partial}{\partial z}, \tag{2}$$

so that the wave-equation became

$$(\mathrm{H} - \mathrm{E}) = 0, \tag{3}$$

where E is defined as the operator

$$\mathrm{E} = -\frac{h}{2\pi i}\frac{\partial}{\partial t}. \tag{4}$$

To generalise the equation (3) for the relativity effect, we shall assume first that the particle of rest mass m_0 is moving in a free space, *i.e.*, $\mathrm{V} = 0$. The velocity u of the particle is given by $u^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$. If we write $\beta = \frac{u}{c}$, the energy E is given by

$$\mathrm{E} = \frac{m_0c^2}{\sqrt{1 - \beta^2}}. \tag{5}$$

Further, for the momenta we have

$$p_x = \frac{m_0\dot{x}}{\sqrt{1-\beta^2}}, \quad p_y = \frac{m_0\dot{y}}{\sqrt{1-\beta^2}}, \quad p_z = \frac{m_0\dot{z}}{\sqrt{1-\beta^2}};$$

hence

$$p_x^2 + p_y^2 + p_z^2 = \frac{m_0^2u^2}{1-\beta^2} = \frac{m_0^2c^2\beta^2}{1-\beta^2} = m_0^2c^2\left(\frac{1}{1-\beta^2} - 1\right),$$

which, on account of (5) gives

$$p_x^2 + p_y^2 + p_z^2 = \frac{E^2}{c^2} - m_0^2c^2,$$

or

$$p_x^2 + p_y^2 + p_z^2 + m_0 c^2 - \frac{E^2}{c^2} = 0.^* \tag{6}$$

The relativistic wave-equation would therefore be

$$\left(p_x^2 + p_y^2 + p_z^2 + m_0^2 c^2 - \frac{E^2}{c^2}\right)\psi = 0, \tag{7}$$

or on substituting from (2) and (4)

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{4\pi^2}{h^2}\, m_0^2c^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\psi = 0. \tag{8}$$

We write $p_t = \frac{E}{c}$ and the wave-equation (7) for a free electron can be written as

$$(-p_t^2 + p_x^2 + p_y^2 + p_z^2 + m_0^2 c^2)\psi = 0. \tag{9}$$

Now consider an external electromagnetic field of scalar potential A_t, and vector potential $\vec{A}$ with components A_x, A_y, A_z, and let the moving particle be an electron of charge $-e$. Then if we write $p_t = \frac{E}{c}$, the equation (15) of § 11·1 (4) becomes

$$-\left(p_t + \frac{e}{c}A_t\right)^2 + \left(p_x + \frac{e}{c}A_x\right)^2 + \left(p_y + \frac{e}{c}A_y\right)^2 + \left(p_z + \frac{e}{c}A_z\right)^2 + m_0^2 c^2 = 0,$$

or, on setting

$$g_t = p_t + \frac{e}{c}A_t, \quad g_x = p_x + \frac{e}{c}A_x, \quad g_y = p_y + \frac{e}{c}A_y, \quad g_z = p_z + \frac{e}{c}A_z, \tag{10}$$

it becomes:

$$-g_t^2 + g_x^2 + g_y^2 + g_z^2 + m_0^2 c^2 = 0. \tag{11}$$

---

[*] This could be obtained immediately from (15) § 11·1 (4) by setting $\phi = A_x = A_y = A_z = 0$.

This means that in the case of an external electromagnetic field, the p's have to be replaced by the g's. This, if we write

$$F \equiv -g_t^2 + g_x^2 + g_y^2 + g_z^2 + m_0^2 c^2,$$

the relativistic wave-equation becomes

$$F\psi \equiv (-g_t^2 + g_x^2 + g_y^2 + g_z^2 + m_0^2 c^2)\,\psi = 0, \tag{12}$$

where ψ is a function of x, y, z, t.

When the relativistic wave-equations (9) and (12) were applied to atomic problems, it was found that the number of stationary states obtained from these equations was half the number required by experiments. Pauli then pointed out that to remove this deficiency, one must introduce the idea of the spinning electron. According to this hypothesis an electron possesses a spin angular momentum $\frac{1}{2}\frac{h}{2\pi}$, and a magnetic moment $\frac{eh}{4\pi m_0 c}$. Pauli and Darwin fitted this hypothesis into the relativistic wave-equation, and showed that the above-mentioned difficulty of the wrong number of stationary states disappeared if account was taken of the electron-spin.

Early in 1928 Dirac pointed out that the spin theory was open to several objections. At the very outset, it is obvious that the idea of a spinning electron is a completely arbitrary assumption. In Dirac's own words, it remains incomprehensible " why Nature should have chosen this particular model for the electron instead of being satisfied with the point charge ". Then again, on the spinning electron model, we should expect that if the electron is moving in a central field of force, the magnitude of its resultant orbital angular momentum should be constant. But, it turns out that this magnitude is not constant, and the model therefore fails in this respect.

Another, and a very serious, objection to the theory of the present section is that the relativistic wave-equations (9) and (12) are non-linear in E or $\frac{\partial}{\partial t}$. The non-relativistic wave-equation (3) is linear in $\frac{\partial}{\partial t}$, so that if the wave-function ψ_0 is known at time $t = 0$, its value at any time t is determined in terms of ψ_0. This property does not hold for the equations (9) and (12) on account of their being non-linear in $\frac{\partial}{\partial t}$.

Finally, it should be remarked that Pauli's spin theory was successful only for those electrons whose velocities were not too great.

A reformulation of the relativistic wave-theory was given by Dirac early in 1928. He showed that if the wave-equation is set up so as to satisfy the requirements of the principles of relativity and quantum mechanics, the existence of the angular momentum and magnetic moment of the electron can be deduced without any arbitrary assumptions about the spin of the electron.

11·3. *Dirac's Relativistic Theory for a Free Electron.*

Dirac looked for a wave-equation which should be invariant under a Lorentz transformation, and should be equivalent to the non-relativistic equation (8) § 11·2. As pointed out in the last section, this equation should be linear in $\partial/\partial t$, *i.e.*, in p_t, and therefore on account of symmetry, it should be linear in p_x, p_y, p_z also.

Dirac assumed therefore the relativistic wave-equation to be of the form

$$(p_t + a_1 p_x + a_2 p_y + a_3 p_z + a_4 m_0 c)\,\psi = 0, \tag{1}$$

where a_1, a_2, a_3, a_4 are dynamical variables or operators which are independent of the p's, thus commuting with t, x, y, z. Moreover, for a particle moving in empty space, all the points in space are equivalent, and therefore the Hamiltonian cannot involve t, x, y, z. The a's are therefore independent of t, x, y, z, and consequently commute with the p's. This means that a_1, a_2, a_3, a_4 are functions of other dynamical variables and not of the co-ordinates and momenta of the electron, and that the wave-function ψ is a function also of these other variables besides t, x, y, z. These other variables are the so-called " spin matrices " or " spinors," which we shall now investigate.

Multiplying (1) by the operator $(-p_t + a_1 p_x + a_2 p_y + a_3 p_z + a_4 m_0 c)$ on the left, we get

$$(-p_t + a_1 p_x + a_2 p_y + a_3 p_z + a_4 m_0 c)\,(p_t + a_1 p_x + a_2 p_y + a_3 p_z + a_4 m_0 c)\,\psi = 0,$$

or

$$\{-p_t^2 + (a_1^2 p_x^2 + a_2^2 p_y^2 + a_3^2 p_z^2) + (a_1 a_2 + a_2 a_1)\,p_x p_y + (a_2 a_3 + a_3 a_2)\,p_y p_z + (a_3 a_1 + a_1 a_3)\,p_z p_x + a_4^2 m_0^2 c^2 + (a_1 a_4 + a_4 a_1)\,m_0 c p_x + (a_2 a_4 + a_4 a_2)\,m_0 c p_y + (a_3 a_4 + a_4 a_3)\,m_0 c p_z\}\,\psi = 0. \tag{2}$$

This equation agrees with the equation (9) § 11·2 if the α's satisfy the following relations

$$\alpha_\mu^2 = 1, \qquad \alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu = 0 \quad (\mu \neq \nu), \qquad (\mu, \nu = 1, 2, 3, 4). \tag{3}$$

The problem now reduces to that of finding four dynamical variables α_μ which satisfy the relations (3). In quantum mechanics, these dynamical variables are matrices. Dirac obtains them with the help of six other matrices σ_1, σ_2, σ_3, ρ_1, ρ_2, ρ_3, which are of the fourth order, and which are as follows :—

$$\sigma_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix},$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \rho_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$\rho_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \qquad \rho_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

It can be easily verified that the σ's and ρ's satisfy the following relations for all r, s, t,$= 1, 2, 3$:—

$$\sigma_r^2 = 1, \qquad \sigma_r\,\sigma_s + \sigma_s\,\sigma_r = 0, \qquad (r \neq s), \tag{4}$$

$$\rho_r^2 = 1, \qquad \rho_r\,\rho_s + \rho_s\,\rho_r = 0, \qquad (r \neq s), \tag{5}$$

$$\rho_r\,\sigma_t = \sigma_t\,\rho_r. \tag{6}$$

The α's are now defined by the following equations* :—

$$\alpha_1 = \rho_1\sigma_1, \quad \alpha_2 = \rho_1\sigma_2, \quad \alpha_3 = \rho_1\sigma_3, \quad \alpha_4 = \rho_3. \tag{7}$$

Obviously, the α's so defined satisfy the relations (3) in virtue of the properties (4), (5) and (6).

The wave-equation (1) now takes the form :

$$\{p_t + \rho_1\,(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z) + \rho_3 m_0 c)\,\psi = 0. \tag{8}$$

This is Dirac's wave-equation for a free electron.

---

* The matrices for the α's are given in § 11·7.

11 · 4. *The Electron under an Arbitrary Field.*

Now suppose that the electron is moving in an external electromagnetic field of scalar potential A_t and vector potential $\vec{A}$. We have shown in § 11·2 that in order to get the Hamiltonian for this case, we must substitute $p_t + \frac{e}{c} A_t$ for p_t, $p_x + \frac{e}{c} A_x$, for p_x, etc., in the Hamiltonian for the case of a free electron. Carrying out this substitution, and remembering the definitions of the g's [(10) § 11 · 2], we get from (8) § 11 · 3 :

$$\{g_t + \rho_1(\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z) + \rho_3 m_0 c\}\,\psi = 0. \tag{1}$$

Writing down the values of the operators g_t, g_x, g_y, g_z, we get Dirac's relativistic wave-equation :

$$\Big[\Big(-\frac{h}{2\pi i c}\frac{\partial}{\partial t} + \frac{e}{c}A_t\Big) + \rho_1\Big\{\sigma_1\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}A_x\Big) + {} + \sigma_2\Big(\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{e}{c}A_y\Big) + \sigma_3\Big(\frac{h}{2\pi i}\frac{\partial}{\partial z} + \frac{e}{c}A_z\Big)\Big\} + \rho_3\, m_0 c\Big]\psi = 0, \tag{2}$$

or

$$\Big\{\Big(-\frac{h}{2\pi i c}\frac{\partial}{\partial t} + \frac{e}{c}A_t\Big) + \alpha_1\Big(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}A_x\Big) + \alpha_2\Big(\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{e}{c}A_y\Big) + \alpha_3\Big(\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{e}{c}A_z\Big) + \alpha_4\, m_0 c\Big\}\,\psi = 0. \tag{3}$$

For the purpose of this section only, suppose that we take a representation of the α's in which all the elements of the matrices representing α_1, α_2, α_3 are real, and all the elements of the matrix representing α_4 are pure imaginary. This is, for instance, possible if in (7) § 11 · 3 we take $\alpha_2 = \rho_3$ and $\alpha_4 = \rho_1\sigma_2$. Then, if we change the sign of i in (3), we see that the complex conjugate of (3) also holds, *viz.*,

$$\Big\{\Big(\frac{h}{2\pi i c}\frac{\partial}{\partial t} + \frac{e}{c}A_t\Big) + \alpha_1\Big(-\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}A_x\Big) + \alpha_2\Big(-\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{e}{c}A_y\Big) + \alpha_3\Big(-\frac{h}{2\pi i}\frac{\partial}{\partial z} + \frac{e}{c}A_z\Big) - \alpha_4\, m_0 c\Big\}\,\psi = 0. \tag{4}$$

Now, if we had an electron with a positive charge $+e$, the Equation (4) would become :

$$\Big\{\Big(\frac{h}{2\pi i c}\frac{\partial}{\partial t} - \frac{e}{c}A_t\Big) - \alpha_1\Big(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}A_x\Big) - \alpha_2\Big(\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{e}{c}A_y\Big) - \alpha_3\Big(\frac{h}{2\pi i}\frac{\partial}{\partial z} + \frac{e}{c}A_z\Big) - \alpha_4\, m_0 c\Big\}\,\psi = 0. \tag{5}$$

This is the same equation as (3). It is obtained from (3) by multiplying both sides by -1. Thus we see that the wave-equation (2) refers to a negative electron (with charge $-e$) as well as to a "positive electron" (with charge $+e$). Half the solution of (2) would refer to the negative electron, and half to the positive electron.

On account of the matrices α being of the fourth order, the wave-function ψ will have four components, two of which will correspond to the negative electron having states of positive energy, and the other two would correspond to the positive electron having states of negative energy. At the time when Dirac first published his theory, the postive electron, or the "positron", was not known to exist. It is generally recognised that with his relativistic electron theory Dirac had predicted the existence of these particles. As remarked in § 3 · 4, the existence of the positron was experimentally demonstrated in 1932 by Carl Anderson.

Dirac has also proved that his relativistic wave-equations (8) § 11 · 3 and (1) § 11 · 4 are invariant under a Lorentz-transformation.

11·5. *Existence of the Magnetic Moment of the Electron.*

To find out how Dirac's wave-equation (1) § 11·4 differs from the non-relativistic wave-equation (10) § 11·2, we multiply the former by a factor, just as we did in § 11·3 to get equation (2) from (1). Remembering the definitions of the g's, *viz.*,

$$g_t = p_t + \frac{e}{c}A_t,\ g_x = p_x + \frac{e}{c}A_x,\ g_y = p_y + \frac{e}{c}A_y,\ g_z = p_z + \frac{e}{c}A_z, \tag{1}$$

Dirac's equation is written in the form

$$\{g_t + \rho_1(\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z) + \rho_3 m_0 c\}\psi = 0. \tag{2}$$

Multiplying (2) on the left with the factor

$$-g_t + \rho_1(\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z) + \rho_3 m_0 c,$$

we obtain

$$\{-g_t + \rho_1(\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z) + \rho_3 m_0 c\}\{g_t + \rho_1(\sigma_1 g_x + \sigma_2 g_y + \\ + \sigma_3 g_z) + \rho_3 m_0 c)\psi = 0,$$

or, on account of (4), (5), (6), § 11·3,

$$\{-g_t^2 + (\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z)^2 + m_0^2c^2 + \rho_1(\sigma_1 g_x g_t - g_t\sigma_1 g_x \\ + \sigma_2 g_y g_t - g_t\sigma_2 g_y + \sigma_3 g_z g_t - g_t\sigma_3 g_z)\}\psi = 0,$$

or, since the σ's commute with the g's:

$$\{-g_t^2 + (\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z)^2 + m_0^2c^2 + \rho_1[\sigma_1(g_xg_t - g_tg_x) + \sigma_2(g_yg_t - g_tg_y) + \sigma_3(g_zg_t - g_tg_z)]\}\psi = 0. \tag{3}$$

Now

$$(\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z)^2 = (\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z)(\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z)$$
$$= (\sigma_1^2 g_x^2 + \sigma_2^2 g_y^2 + \sigma_3^2 g_z^2) + [(\sigma_1\sigma_2 g_x g_y + \sigma_2\sigma_1 g_y g_x) + (\sigma_2\sigma_3 g_y g_z + \sigma_3\sigma_2 g_z g_y) + (\sigma_3\sigma_1 g_z g_x + \sigma_1\sigma_3 g_x g_z)].$$

But from the definitions of the σ matrices, it can be easily proved that

$$\sigma_1\sigma_2 = i\sigma_3 = -\sigma_2\sigma_1, \quad \sigma_2\sigma_3 = i\sigma_1 = -\sigma_3\sigma_2, \quad \sigma_3\sigma_1 = i\sigma_2 = -\sigma_1\sigma_3,$$

and as shown in (4) § 11·3 (4)

$$\sigma_1^2 = 1, \quad \sigma_2^2 = 1, \quad \sigma_3^2 = 1.$$

We have therefore

$$(\sigma_1 g_x + \sigma_2 g_y + \sigma_3 g_z)^2 = (g_x^2 + g_y^2 + g_z^2) + i\{\sigma_3(g_xg_y - g_yg_x) + \sigma_1(g_yg_z - g_zg_y) + \sigma_2(g_zg_x - g_xg_z)\}. \tag{5}$$

Now consider

$$\Big(g_xg_y - g_yg_x\Big)\psi = \Big\{\Big(p_x + \frac{e}{c}A_x\Big)\Big(p_y + \frac{e}{c}A_y\Big) - \Big(p_y + \frac{e}{c}A_y\Big)\Big(p_x + \frac{e}{c}A_x\Big)\Big\}\psi$$

$$= \Big\{\Big(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}A_x\Big)\Big(\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{e}{c}A_y\Big) - \Big(\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{e}{c}A_y\Big)\Big(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}A_x\Big)\Big\}\psi$$

$$= \Big(\frac{h}{2\pi i}\Big)^2\frac{\partial^2\psi}{\partial x\partial y} + \frac{h}{2\pi i}\frac{\partial}{\partial x}\Big(\frac{e}{c}A_y\psi\Big) - \frac{e}{c}A_x\frac{h}{2\pi i}\frac{\partial\psi}{\partial y} + \frac{c^2}{c^2}A_xA_y\psi$$
$$- \Big(\frac{h}{2\pi i}\Big)^2\frac{\partial^2\psi}{\partial y\partial x} - \frac{h}{2\pi i}\frac{\partial}{\partial y}\Big(\frac{e}{c}A_x\psi\Big) - \frac{e}{c}A_y\frac{h}{2\pi i}\frac{\partial\psi}{\partial x} - \frac{e^2}{c^2}A_yA_x\psi$$

$$= \frac{he}{2\pi ic}\frac{\partial}{\partial x}(A_y\psi) + \frac{he}{2\pi ic}A_x\frac{\partial x}{\partial y} - \frac{he}{2\pi ic}\frac{\partial}{\partial y}(A_x\psi) - \frac{he}{2\pi ic}A_y\frac{\partial\psi}{\partial x}$$
$$= \frac{he}{2\pi ic}\Big(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\Big)\psi$$

$$= \frac{he}{2\pi ic}H_z\psi, \tag{6}$$

because of (1) § 11·1.

Similarly we get

$$(g_y g_z - g_z g_y)\,\psi = \frac{he}{2\pi ic}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\psi = \frac{he}{2\pi ic}\,H_x\,\psi, \tag{6 a}$$

$$(g_z g_x - g_x g_z)\,\psi = \frac{he}{2\pi ic}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\psi = \frac{he}{2\pi ic}\,H_y\,\psi. \tag{6 b}$$

Further, we find

$$\begin{aligned}(g_x g_t - g_t g_x)\,\psi = \Big\{&\left(p_x + \frac{e}{c}\,A_x\right)\left(p_t + \frac{e}{c}\,A_t\right) \\ &- \left(p_t + \frac{e}{c}\,A_t\right)\left(p_x + \frac{e}{c}\,A_x\right)\Big\}\,\psi,\end{aligned}$$

$$\begin{aligned}= \Big\{&\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}\,A_x\right)\left(\frac{-h\,\partial}{2\pi ic\,\partial t} + \frac{e}{c}\,A_t\right) \\ &- \left(\frac{-h\,\partial}{2\pi ic\,\partial t} + \frac{e}{c}\,A_t\right)\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}\,A_x\right)\Big\}\,\psi,\end{aligned}$$

$$\begin{aligned}= &-\left(\frac{h}{2\pi i}\right)^2 \frac{1}{c}\frac{\partial^2\psi}{\partial x\partial t} + \frac{h}{2\pi i}\frac{\partial}{\partial x}\left(\frac{e}{c}\,A_t\,\psi\right) - \frac{e}{c^2}\,A_x\,\frac{h}{2\pi i}\frac{\partial\psi}{\partial t} + \frac{e^2}{c^2}\,A_x A_t\psi \\ &+ \left(\frac{h}{2\pi i}\right)^2 \frac{1}{c}\frac{\partial^2\psi}{\partial x\partial t} + \frac{h}{2\pi ic}\frac{\partial}{\partial t}\left(\frac{e}{c}\,A_x\,\psi\right) - \frac{e}{c}\,A_t\,\frac{h}{2\pi i}\frac{\partial\psi}{\partial x} - \frac{e^2}{c^2}\,A_t A_x\psi,\end{aligned}$$

$$= \frac{he}{2\pi ic}\frac{\partial}{\partial x}\,(A_t\,\psi) - \frac{he}{2\pi ic^2}\,A_x\,\frac{\partial\psi}{\partial t} + \frac{he}{2\pi ic^2}\frac{\partial}{\partial t}\,(A_x\,\psi) - \frac{he}{2\pi ic}\,A_t\,\frac{\partial\psi}{\partial x},$$

$$= \frac{he}{2\pi ic}\left(\frac{\partial A_t}{\partial x} + \frac{1}{c}\frac{\partial A_x}{\partial t}\right)\psi = \frac{-he}{2\pi ic}\,E_x\,\psi, \tag{7}$$

because of (1) § 11·1

Similarly,

$$(g_y g_t - g_t g_y)\,\psi = \frac{he}{2\pi ic}\left(\frac{\partial A_t}{\partial y} + \frac{1}{c}\frac{\partial A_y}{\partial t}\right)\psi = \frac{-he}{2\pi ic}\,E_y\,\psi, \tag{7 a}$$

$$(g_z g_t - g_t g_z)\,\psi = \frac{he}{2\pi ic}\left(\frac{\partial A_t}{\partial z} + \frac{1}{c}\frac{\partial A_z}{\partial t}\right)\psi = \frac{-he}{2\pi ic}\,E_z\,\psi. \tag{7 b}$$

Substituting (5), (6), (7) in (3), we obtain

$$\begin{aligned}\Big\{-g_t^2 + g_1^2 + g_2^2 + g_3^2 + m_0^2c^2 + \frac{he}{2\pi c}\,(\sigma_1\,H_x + \sigma_2 H_y + \sigma_3 H_z) &+ \\ \frac{ihe}{2\pi c}\,\rho_1\,(\sigma_1 E_x + \sigma_2 E_y + \sigma_3 E_z)\Big\}\,\psi = 0&\end{aligned} \tag{8}$$

or, substituting the value of the g's from (1) :

$$\Big\{-\left(p_t+\frac{e}{c}A_t\right)^2+\left(p_x+\frac{e}{c}A_x\right)^2+\left(p_y+\frac{e}{c}A_y\right)^2+$$
$$\left(p_z+\frac{e}{c}A_z\right)^2+m^2c^2+\frac{he}{2\pi c}(\sigma_1H_x+\sigma_2H_y+\sigma_3H_2)$$
$$+\frac{ihe}{2\pi c}\rho_1(\sigma_1E_x+\sigma_2E_y+\sigma_3E_z)\Big\}\psi=0. \tag{9}$$

The Hamiltonian in this equation differs from that in (10) § 11·3 by the terms $\frac{he}{2\pi c}(\sigma_1H_x+\sigma_2H_y+\sigma_3H_z)+\frac{ihe}{2\pi c}\rho_1(\sigma_1E_x+\sigma_2E_y+\sigma_3E_z)$. These two terms, divided by 2 m_0, can be considered as an increase in the potential energy of the electron due to its spin. Thus the additional potential energy is

$$\frac{1}{2m_0}\cdot\frac{he}{2\pi c}(\sigma_1H_x+\sigma_2H_y+\sigma_3H_2)$$
$$+\frac{1}{2m_0}\frac{he}{2\pi c}i\rho_1(\sigma_1E_x+\sigma_2E_y+\sigma_3E_z). \tag{10}$$

The electron will therefore behave as though it has a magnetic moment with components

$$\left(\frac{eh}{4\pi m_0c}\sigma_1,\ \frac{eh}{4\pi m_0c}\sigma_2,\ \frac{eh}{4\pi m_0c}\sigma_3\right), \tag{11}$$

and an electric moment with components

$$\left(\frac{eh}{4\pi m_0c}i\rho_1\sigma_1,\ \frac{eh}{4\pi m_0c}i\rho_1\sigma_2,\ \frac{eh}{4\pi m_0c}i\rho_1\sigma_3\right). \tag{12}$$

The magnetic moment has, therefore, just the amount assumed in the spinning electron theory, § 11·2, and which is here derived, according to Dirac, as a natural consequence of the relativistic wave-equation.

The electric moment, however, is a pure imaginary quantity, and cannot be considered to have any physical meaning. It has arisen only after multiplying the real Hamiltonian in (1) § 11·5 with a factor in order to compare it with the classical Hamiltonian in (10) § 11·2. This being rather an artificial operation, we must not be surprised if it leads to some terms which have no physical meaning.

11·6. *Existence of the Spin Angular Momentum of the Electron.*

We remarked at the end of § 11·2 that the spinning electron theory postulated the existence of an intrinsic (spin) angular momentum of the electron of amount $\frac{1}{2}\frac{h}{2\pi}$. We have now to prove the existence of this spin angular momentum on Dirac's theory. Since this angular momentum does not give rise to any potential energy, it did not appear in the result of calculations of § 11·5. To demonstrate its existence, we shall consider the special case of the motion of an electron in a central field of force, and shall determine the angular momentum integrals.

Take the centre of force as origin, rectangular Cartesian co-ordinates x, y, z and corresponding momenta p_x, p_y, p_z. μ is the mass of the electron, and V the potential energy which is a function of $x^2 + y^2 + z^2$.

The classical Hamiltonian, neglecting the relativity effect, is

$$\mathrm{H} = \frac{1}{2\mu}(p_x^2 + p_y^2 + p_z^2) + \mathrm{V}. \tag{1}$$

If $\vec{m}$ is the orbital angular momentum of the electron, we proved in § 6·8 (1) that each of m_x, m_y, m_z commutes with any function of $x^2 + y^2 + z^2$ and $p_x^2 + p_y^2 + p_z^2$. The Hamiltonian H, as given in (1), being a function of $x^2 + y^2 + z^2$ and $p_x^2 + p_y^2 + p_z^2$, we see that each of m_x, m_y, m_z, commutes with H. We conclude that the orbital angular momentum is a constant of the motion, just as in Newtonian dynamics.

However, if we take account of the relativity effect also, it turns out that the orbital angular momentum $\vec{m}$ alone is not a constant of the motion. To get such a constant, we have to add another term which comes out to be just the required spin angular momentum.

The force on the electron is given by the potential V (r) which can be taken to be the scalar potential A_t, the vector potential $\vec{\mathrm{A}} \equiv (\mathrm{A}_x, \mathrm{A}_y, \mathrm{A}_z)$ being zero. Thus

$$\mathrm{A}_x = \mathrm{A}_y = \mathrm{A}_z = 0, \ \mathrm{A}_t = \mathrm{V}(r). \tag{1}$$

From Dirac's wave-equation (1) § 11·4, we get then

$$\mathrm{F}\psi \equiv \left\{p_t + \frac{e}{c}\mathrm{V} + \rho_1(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z) + \rho_3 m_0 c\right\}\psi = 0. \tag{2}$$

To get the Hamiltonian H, we remark that $p_t = \frac{\mathrm{E}}{c}$, where E is the energy parameter. (Since we require only periodic solutions of the wave-equation $F\psi = 0$, p_t is no more considered as an operator).

The equation (2) can therefore be written as

$$\left\{\frac{E}{c} - \left[-\frac{e}{c}V - \rho_1(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z) - \rho_3 m_0 c\right]\right\}\psi = 0,$$

or, on multiplying throughout by c,

$$\{E - [-eV - c\rho_1(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z) - \rho_3 m_0 c^2]\}\psi = 0. \tag{3}$$

Thus we see that the relativistic Hamiltonian is

$$H = -eV - c\rho_1(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z) - \rho_3 m_0 c^2. \tag{4}$$

The classical Hamiltonian would have been simply $-eV$, so that the last two terms in (4) are contributed by the relativity-effect.

We shall prove now that the orbital angular momentum $\vec{m}$ is not a constant of the motion, *i.e.*, the components m_x, m_y, m_z as defined in (1) § 6·8 (1) do not commute with H given by (4).

For

$$\begin{aligned} m_x H - H m_x &= -c\rho_1\ \{m_x(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z) - (\sigma_1 p_x + \sigma_2 p_y \\ &\qquad + \sigma_3 p_z)m_x\} \\ &= -c\rho_1\ \{\sigma_1(m_x p_x - p_x m_x) + \sigma_2(m_x p_y - p_y m_x) \\ &\qquad + \sigma_3(m_x p_z - p_z m_x)\} \\ &= -\frac{ih}{2\pi}c\rho_1\{\sigma_2 p_z - \sigma_3 p_y\}\psi \neq 0, \end{aligned} \tag{5}$$

on account of (5) § 6·8 (1). It can be similarly proved that m_y and m_z do not commute with H. Thus the orbital angular momentum is not a constant of the motion.

We have further from (4)

$$\begin{aligned} \frac{1}{2}\frac{h}{2\pi}\sigma_1 H - H\frac{1}{2}\frac{h}{2\pi}\sigma_1 &= -\frac{hc}{4\pi}\rho_1\{\sigma_1(\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z) \\ &\qquad - (\sigma_1 p_x + \sigma_2 p_y + \sigma_3 p_z)\sigma_1\} \\ &= -\frac{hc}{4\pi}\rho_1\left\{(\sigma_1\sigma_2 - \sigma_2\sigma_1)p_y + (\sigma_2\sigma_3 - \sigma_3\sigma_1)p_z\right\} \\ &= -\frac{hc}{4\pi}\rho_1\left\{2i\sigma_3 p_y - 2i\sigma_2 p_z\right\}, \end{aligned}$$

on account of (4) § 11·5. Thus we get

$$\frac{1}{2}\frac{h}{2\pi}\sigma_1 H - H\frac{1}{2}\frac{h}{2\pi}\sigma_1 = -\frac{ih}{2\pi}c\rho_1\left\{\sigma_3 p_y - \sigma_2 p_z\right\}. \tag{6}$$

From (5) and (6) we find therefore that

$$\left(m_x + \frac{1}{2}\frac{h}{2\pi}\sigma_1\right)\mathrm{H} - \mathrm{H}\left(m_x + \frac{1}{2}\frac{h}{2\pi}\sigma_1\right) = 0, \qquad (7)$$

and similarly, that $m_y + \frac{1}{2}\frac{h}{2\pi}\sigma_2$ and $m_z + \frac{1}{2}\frac{h}{2\pi}\sigma_3$ commute with H. Thus if we define $\vec{\mathrm{M}}$ by the relations

$$\mathrm{M}_x = m_x + \frac{1}{2}\frac{h}{2\pi}\sigma_1,\ \mathrm{M}_y = m_y + \frac{1}{2}\frac{h}{2\pi}\sigma_2,\ \mathrm{M}_z = m_z + \frac{1}{2}\frac{h}{2\pi}\sigma_3, \qquad (8)$$

we see that $\vec{\mathrm{M}}$, and not the orbital angular momentum $\vec{m}$ alone, is a constant of the motion.

This result can be interpreted to mean that the electron has a spin angular momentum with components $\left(\frac{1}{2}\frac{h}{2\pi}\sigma_1,\ \frac{1}{2}\frac{h}{2\pi}\sigma_3,\ \frac{1}{2}\frac{h}{2\pi}\sigma_3\right)$, which is just the spin angular momentum postulated in the spinning electron theory of § 11·2.

11·7. *The Theory of the Hydrogen Atom; Fine Structure of Spectral Lines.*

As a particular example of Dirac's theory, we shall solve the problem of the hydrogen-atom, obtaining the fine structure of the lines observed in instruments of high resolving power. Sommerfeld's treatment of this problem on Bohr's theory has been described in § 5·7. Dirac solved the problem as an example of his theory, but immediately afterwards C. G. Darwin gave a solution which is simpler, and which is followed here.

For the hydrogen atom the electron is under the central field from the nucleus only, there being no magnetic field. Thus, if r is the distance of the electron from the nucleus:

$$\mathrm{A}_t = \mathrm{V} = \frac{e}{r},\ \mathrm{A}_x = \mathrm{A}_y = \mathrm{A}_z = 0. \qquad (1)$$

Dirac's equation (1) § 11·4 therefore becomes on account of (7) § 11·3:

$$\left\{\left(p_0 + \frac{e\mathrm{V}}{c}\right) + \alpha_1 p_x + \alpha_2 p_y + \alpha_3 p_z + \alpha_4 m_0 c\right\}\psi = 0. \qquad (2)$$

From the definition of the α's given in (7) § 11·3, and the matrices for the σ's and ρ's, we get

$$\alpha_1 = \rho_1\sigma_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad \alpha_2 = \rho_1\sigma_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix},$$

$$\alpha_3 = \rho_1\sigma_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \alpha_4 = \rho_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3}$$

According to these 4-rowed matrices, there are four wave-functions ψ_1, ψ_2, ψ_3, ψ_4 which can be written in a column

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix}. \tag{4}$$

Since the field is central and we want periodic solutions, we write for p_t, the momentum conjugate to the energy E, its value $\frac{E}{c}$, and for p_x the operator $\frac{h}{2\pi i}\frac{\partial}{\partial x}$, etc.

Thus corresponding to the four rows of the matrices (3), we get from (2) the following four equations for the four wave-functions:

$$\frac{2\pi i}{h}\left(\frac{E + eV}{c} + m_0c\right)\psi_1 + \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\psi_4 + \frac{\partial}{\partial z}\psi_3 = 0, \tag{5}$$

$$\frac{2\pi i}{h}\left(\frac{E + eV}{c} + m_0c\right)\psi_2 + \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\psi_3 - \frac{\partial}{\partial z}\psi_4 = 0,$$

$$\frac{2\pi i}{h}\left(\frac{E + eV}{c} - m_0c\right)\psi_3 + \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\psi_2 + \frac{\partial}{\partial z}\psi_1 = 0,$$

$$\frac{2\pi i}{h}\left(\frac{E + eV}{c} - m_0c\right)\psi_4 + \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\psi_1 - \frac{\partial}{\partial z}\psi_2 = 0.$$

Equations (5) show that the two functions (ψ_1, ψ_2) form one set of solutions, and the other two functions (ψ_3, ψ_4) form another set.

As the problem is that of a radial force, we must transform to polar co-ordinates (r, θ, ϕ) given by

$$x = r \sin\theta \cos\phi, \quad y = r \sin\theta \sin\phi, \quad z = r \cos\theta. \tag{6}$$

Moreover, since the potential V is a function of r only, being independent of θ and ϕ, the solutions of (5), *i.e.*, the four wave-functions must involve the spherical harmonics S_k^m:

$$\mathrm{S}_k^m = \mathrm{P}_k^m(\theta)\ e^{im\phi}, \tag{7}$$

where P_k^m are the associated Legendre functions given by

$$\mathrm{P}_k{}^m = (k - m)\,!\,\sin^m\theta \left(\frac{d}{d\cos\theta}\right)^{k+m} \frac{(\cos^2\theta - 1)^k}{2^k\,k\,!}, \tag{8}$$

k being any positive integer, and m being any integer between $-k$ and $+k$ inclusive.

Now if F (r) is any function of r only, then substituting the values of x, y, z, from (6) it is easily calculated that

$$\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\mathrm{F}(r)\,\mathrm{S}_k^m = \frac{1}{2k+1} \times$$

$$\left\{\left(\frac{d\mathrm{F}}{dr} - \frac{k}{r}\mathrm{F}\right)\mathrm{S}_{k+1}^{m+1} - (k-m)(k-m-1)\left(\frac{d\mathrm{F}}{dr} + \frac{k+1}{r}\mathrm{F}\right)\mathrm{S}_{k-1}^{m+1}\right\},$$

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\mathrm{F}(r)\,\mathrm{S}_k^m = \frac{1}{2k+1} \times$$

$$\left\{-\left(\frac{d\mathrm{F}}{dr} - \frac{k}{r}\mathrm{F}\right)\mathrm{S}_{k+2}^{m-1} + (k+m)(k+m-1)\left(\frac{d\mathrm{F}}{dr} + \frac{k+1}{r}\mathrm{F}\right)\mathrm{S}_{k-1}^{m-1}\right\},$$

$$\frac{\partial}{\partial z}\mathrm{F}(r)\,\mathrm{S}_k^m = \frac{1}{2k+1} \times$$

$$\left\{\left(\frac{d\mathrm{F}}{dr} - \frac{k}{r}\mathrm{F}\right)\mathrm{S}_{k+1}^m + (k+m)(k-m)\left(\frac{d\mathrm{F}}{dr} + \frac{k+1}{r}\mathrm{F}\right)\mathrm{S}_{k-1}^m\right\}. \tag{9}$$

Thus considering (5) and (9), we find that if we want each one of the ψ's to involve a single spherical harmonic only, we must write for a trial solution:

$$\begin{aligned} \psi_1 &= -ia_1 f(r)\,\mathrm{S}_{k+1}^m, & \psi_2 &= -ia_2 f(r)\,\mathrm{S}_{k+1}^{m+1}, \\ \psi_3 &= a_3\, g(r)\,\mathrm{S}_k^m, & \psi_4 &= a_4\, g(r)\,\mathrm{S}_k^{m+1}. \end{aligned} \tag{10}$$

The same radial function $f(r)$ has to be taken for the two wave-functions ψ_1 and ψ_2 of the first set, and similarly, the same function $g(r)$ has to be taken for ψ_3 and ψ_4. Moreover, $-i$ is introduced in ψ_1 and ψ_2 to make f real. The constants a_1, a_2, a_3, a_4 are introduced so that all the four equations (5) may be satisfied. Thus substituting (10) in (5), and taking account of the relations (9), we obtain

$$\left.\begin{aligned}
&\frac{2\pi}{h}\left(\frac{E+eV}{c}+m_0c\right)a_1 f\,S_{k+1}^{m}+\frac{a_4}{2k+1}\times\\
&\Big\{-\Big(\frac{dg}{dr}-\frac{k}{r}\Big)S_{k+1}^{m}+(k+m+1)(k+m)\Big(\frac{dg}{dr}+\frac{k+1}{r}\Big)S_{k-1}^{m}\Big\}\\
&\qquad+\frac{a_3}{2k+1}\Big\{\Big(\frac{dg}{dr}-\frac{k}{r}g\Big)S_{k+1}^{m}+(k+m)(k-m)\times\\
&\qquad\qquad\qquad\Big(\frac{dg}{dr}+\frac{k+1}{r}g\Big)S_{k-1}^{m}\Big\}=0,\\
&\frac{2\pi}{h}\left(\frac{E+eV}{c}+m_0c\right)a_2 f\,S_{k+1}^{m+1}+\frac{a_3}{2k+1}\times\\
&\Big\{\Big(\frac{dg}{dr}-\frac{k}{r}g\Big)S_{k+1}^{m+1}-(k-m)(k-m-1)\Big(\frac{dg}{dr}+\frac{k+1}{r}g\Big)S_{k-1}^{m+1}\Big\}\\
&\qquad-\frac{a_4}{2k+1}\Big\{\Big(\frac{dg}{dr}-\frac{kg}{r}\Big)S_{k+1}^{m+1}+(k+m+1)(k-m-1)\times\\
&\qquad\qquad\qquad\Big(\frac{dg}{dr}+\frac{k+1}{r}g\Big)S_{k-1}^{m+1}\Big\}=0
\end{aligned}\right\}\qquad(11)$$

$$\left.\begin{aligned}
&\frac{2\pi}{h}\left(\frac{E+eV}{c}-m_0c\right)a_3 g\,S_{k}^{m}-\frac{a_2}{2k+3}\times\\
&\Big\{-\Big(\frac{df}{dr}-\frac{k+1}{r}f\Big)S_{k+2}^{m}+(k+m+2)(k+m+1)\Big(\frac{df}{dr}+\frac{k+2}{r}f\Big)S_{k}^{m}\Big\}\\
&\qquad-\frac{a_1}{2k+3}\Big\{\Big(\frac{df}{dr}-\frac{k+1}{r}f\Big)S_{k+2}^{m}+(k+m+1)(k-m+1)\\
&\qquad\qquad\qquad\times\Big(\frac{df}{dr}+\frac{k+2}{r}f\Big)S_{k}^{m}\Big\}=0,\\
&\frac{2\pi}{h}\left(\frac{E+eV}{c}-m_0c\right)a_4 g\,S_{k}^{m+1}-\frac{a_1}{2k+3}\Big\{\Big(\frac{df}{dr}-\frac{k+1}{r}f\Big)S_{k+2}^{m+1}\\
&\qquad-(k-m+1)(k-m)\Big(\frac{df}{dr}+\frac{k+2}{r}f\Big)S_{k}^{m+1}\Big\}\\
&\qquad+\frac{a_2}{2k+3}\Big\{\Big(\frac{df}{dr}-\frac{k+1}{r}f\Big)S_{k+2}^{m+1}+(k+m+2)(k-m)\\
&\qquad\qquad\qquad\Big(\frac{df}{dr}+\frac{k+2}{r}f\Big)S_{k}^{m+1}\Big\}=0.
\end{aligned}\right\}\qquad(12)$$

So far the four constants a_1, etc. have been quite arbitrary. Now we choose them so that the four equations (11) and (12) become as simple as possible. This will be the case if out of the two spherical harmonics in each equation one disappears, and the remaining equations are consistent.

The two equations (11) show that if we choose

$$a_4 (k + m + 1) + a_3 (k - m) = 0, \tag{13}$$

then the term in S_{k-1}^{m} from the first equation, and that in S_{k+1}^{m+1} from the second equation will disappear.

Similarly, the two equations (12) show that if we choose

$$a_1 = a_2, \tag{14}$$

then the term in S_{k+2}^{m} from the first equation, and that in S_{k+2}^{m+1} from the second equation will disappear.

The relations (13) and (14) will be satisfied if we write definitely :

$$a_1 = 1, \quad a_2 = 1, \quad a_3 = k + m + 1, \quad a_4 = -k + m, \tag{15}$$

because the ratio $a_1 : a_4$ can be incorporated in $f : g$.

Substituting (15) in (11) and (12), and equating the coefficient of the remaining spherical harmonic to zero, we find that the two equations (11) reduce to the same equation

$$\frac{2\pi}{h}\left(\frac{\mathrm{E} + e\mathrm{V}}{c} + m_0 c\right) f + \frac{dg}{dr} - \frac{k}{r} g = 0 ; \tag{16}$$

and the two equations (12) reduce to the same equation

$$-\frac{2\pi}{h}\left(\frac{\mathrm{E} + e\mathrm{V}}{c} - m_0 c\right) g + \frac{df}{dr} + \frac{k + 2}{r} f = 0. \tag{17}$$

(16) and (17) form a system of two ordinary linear differential equations for the determination of the two unknown functions f and g.

Writing the value of $\mathrm{V} = \frac{e}{r}$, and setting for brevity :

$$a^2 = \frac{2\pi}{h}\left(\frac{\mathrm{E}}{c} + m_0 c\right), \quad b^2 = \frac{2\pi}{h}\left(-\frac{\mathrm{E}}{c} + m_0 c\right), \quad \alpha = \frac{2\pi e^2}{ch}, \tag{18}$$

where α is Sommerfeld's fine structure constant (21) § 5·7, the two equations (16) and (17) become :

$$\left(a^2 + \frac{\alpha}{r}\right) f + \left(\frac{d}{dr} - \frac{k}{r}\right) g = 0, \tag{19}$$

$$\left(\frac{d}{dr} + \frac{k + 2}{r}\right) f + \left(b^2 - \frac{\alpha}{r}\right) g = 0.$$

We transform the dependent variables f, g to f', g' by writing

$$f = e^{-\lambda r} f', \; g = e^{-\lambda r} g', \tag{20}$$

where λ is a constant whose value is to be determined.

Substituting (20) in (19), and dividing out by $e^{-\lambda r}$, we get:

$$\begin{aligned} \left(a^2 + \frac{\alpha}{r}\right) f' + \left(\frac{d}{dr} - \lambda - \frac{k}{r}\right) g' &= 0, \\ \left(\frac{d}{dr} - \lambda + \frac{k+2}{r}\right) f' + \left(b^2 - \frac{\alpha}{r}\right) g' &= 0. \end{aligned} \tag{21}$$

We now try for a solution of (21) in which f' and g' are in the form of the series.

$$\begin{aligned} f' &= a_0 r^{\beta} + a_1 r^{\beta - 1} + a_2 r^{\beta - 2} + \cdots, \\ g' &= b_0 r^{\beta} + b_1 r^{\beta - 1} + b_2 r^{\beta - 2} + \cdots, \end{aligned} \tag{22}$$

where the constant β, whose value is to be determined, need not be an integer. We have taken series of decreasing powers for f' and g', because these functions have to be finite at $r = 0$.

Substituting (22) in (21), we obtain:

$$\begin{aligned} & a^2 \{a_0 r^{\beta} + a_1 r^{\beta - 1} + a_2 r^{\beta - 2} + \cdots\} + \\ & \quad \alpha \{a_0 r^{\beta - 1} + a_1 r^{\beta - 2} + a_2 r^{\beta - 3} + \cdots\} + \\ & \quad \{b_0 \beta r^{\beta - 1} + b_1 (\beta - 1) r^{\beta - 2} + \cdots\} - \\ & \quad \lambda \{b_0 r^{\beta} + b_1 r^{\beta - 1} + b_2 r^{\beta - 2} + \cdots\} - \\ & \quad k \{b_0 r^{\beta - 1} + b_1 r^{\beta - 2} + b_2 r^{\beta - 3} + \cdots\} = 0, \\ & \{a_0 \beta r^{\beta - 1} + a_1 (\beta - 1) r^{\beta - 2} + a_2 (\beta - 2) r^{\beta - 3} + \cdots\} - \\ & \quad \lambda \{a_0 r^{\beta} + a_1 r^{\beta - 1} + a_2 r^{\beta - 2} + \cdots\} + \\ & \quad (k + 2) \{a_0 r^{\beta - 1} + a_1 r^{\beta - 2} + \cdots\} + \\ & \quad b^2 \{b_0 r^{\beta} + b_1 r^{\beta - 1} + b_2 r^{\beta - 2} + \cdots\} - \\ & \quad \alpha \{b_0 r^{\beta - 1} + b_1 r^{\beta - 2} + b_2 r^{\beta - 3} + \cdots\} = 0. \end{aligned} \tag{23}$$

We equate the coefficients of successive powers of r in the two equations (23) to zero, and obtain the values of λ, β, $a_0 : b_0$, $a_1 : b_1$, $a_2 : b_2$, etc. Thus, equating the coefficients of r^{β} to zero, we get

$$a^2 a_0 - \lambda b_0 = 0, \quad b^2 b_0 - \lambda a_0 = 0,$$

from which we find

$$\lambda = \frac{a^2 a_0}{b_0} = \frac{b^2 b_0}{a_0}, \; \textit{i.e.,} \; \frac{b_0}{a_0} = \frac{a}{b} \text{ and } \lambda = ab. \tag{24}$$

We have selected the positive sign for λ, because the functions f and g have to remain finite at $r = \infty$.

Next, equating the coefficients of $r^{\beta-1}$ in (23), we get:

$$a^2 a_1 + \alpha a_0 - \lambda b_1 + (\beta - k)\, b_0 = 0,$$
$$b^2 b_1 - \alpha\beta_0 - \lambda a_1 + (\beta + k + 2)\; a_0 = 0.$$

Substituting the values of $\dfrac{b_0}{a_0}$ and λ from (24), and eliminating both a_1, b_1 from these two equations, we find

$$\beta = -1 + \alpha\,\frac{a^2 - b^2}{2ab}\,. \tag{25}$$

Finally, equating the coefficient of $r^{\beta-(s+1)}$ in (23), we obtain:

$$a^2 a_{s+1} + \alpha a_s - \lambda b_{s+1} + (\beta - k - s) b_s = 0,$$
$$b^2 b_{s+1} - \alpha b_s - \lambda a_{s+1} + (\beta + k + 2 - s)\; a_s = 0. \tag{26}$$

Writing the value of $\lambda = ab$, transposing, and multiplying the first equation by b and the second by $-a$, we get:

$$a^2 b a_{s+1} - ab^2 b_{s+1} = -\, b\alpha a_s - b\,(\beta - k - s)\, b_s,$$
$$a^2 b a_{s+1} - ab^2\, b_{s+1} = a\,(\beta + k + 2 - s)\, a_s - a\alpha b_s.$$

Equating the two right-hand sides, and transposing, we obtain

$$a\, a_s \left\{\beta + k + 2 - s + \alpha\,\frac{b}{a}\right\} + b\, b_s \left\{\beta - k - s - \alpha\,\frac{a}{b}\right\} = 0,$$

giving

$$\frac{a_s}{-\dfrac{b}{a}\left\{\beta - k - s - \alpha\,\dfrac{a}{b}\right\}} = \frac{b_s}{\left\{\beta + k + 2 - s + \alpha\,\dfrac{b}{a}\right\}} = c_s \quad \text{(say)},$$

where c_s is a new constant to be determined. Then for all $s \geqslant 1$:

$$\begin{aligned} a_s &= c_s\,\frac{b}{a}\left\{\alpha\,\frac{a}{b} - \beta + k + s)\right\}, \\ b_s &= c_s\left\{\alpha\,\frac{b}{a} + \beta + k + 2 - s\right\}, \end{aligned} \tag{27}$$

Substituting these values in any one of the two equations (26), we find the recurrence formula for c_s:

$$ab\,(2s + 2)\, c_{s+1} = -\, c_s\,\{(s - \beta - 1)^2 - (j^2 - \alpha^2)\}, \tag{28}$$

where we have written

$$j = k + 1. \tag{29}$$

If we write also

$$\gamma^2 = j^2 - \alpha^2. \tag{30}$$

and so $\gamma = \sqrt{j^2 - \alpha^2}$, supposed positive, we have from (28):

$$2a\, b\, (s + 1)\, c_{s+1} = - c_s\, (\beta + 1 - s - \gamma)\, (\beta + 1 - s + \gamma). \tag{31}$$

We can find c_1 direct from the value of a_1 or b_1, and then we can determine c_2, c_3, etc. successively from (31). This gives

$$c_s = \frac{(-1)^s}{2^s . s\,!\,(ab)^s} \Big\{(\beta - \gamma + 1)\,(\beta - \gamma) \cdots (\beta - \gamma - s + 2) \times \\ (\beta + \gamma + 1)\,(\beta + \gamma) \cdots (\beta + \gamma - s + 2)\Big\}. \tag{32}$$

The functions f', g' are therefore of the type:

$$\begin{aligned} f' &= \sum_s c_s \frac{b}{a}\left(\alpha \frac{a}{b} - \beta + k + s\right) r^{\beta - s}, \\ g' &= \sum_s c_s \left(\alpha \frac{b}{a} + \beta + k + 2 - s\right) r^{\beta - s}.. \end{aligned} \tag{33}$$

As remraked above, the boundary condition requires that these functions should remain finite for $r = 0$. The series therefore should terminate at a value of s, such that $\beta - s \geqslant 0$ (otherwise we would have a negative power of r).

Supposing that the series terminates at $s = n$, where n is a positive integer or zero, then c_{n+1} should be zero. On account of (31), this would be the case only when the factor

$$\beta + 1 - n - \gamma = 0, \quad \text{or } \beta = \gamma + n - 1. \tag{34}$$

The two equations (25) an (34) give us the eigen-values of the energy parameter. Thus

$$\begin{aligned} \beta + 1 = \gamma + n &= \alpha \frac{a^2 - b^2}{2ab}, \\ &= \alpha \frac{E}{\sqrt{m_0{}^2 c^4 - E^2}}, \end{aligned} \tag{35}$$

on substituting the values of a, b from (18). Solving this for E, we find

$$\begin{aligned} E &= m_0 c^2 \left\{1 + \frac{\alpha^2}{(n + \gamma)^2}\right\}^{-\frac{1}{2}} \\ &= m_0 c^2 \left\{1 + \frac{\alpha^2}{(n + \sqrt{j^2 - \alpha^2})^2}\right\}^{-\frac{1}{2}}. \end{aligned} \tag{36}$$

This is the same formula as that given by Sommerfeld in (21 *a*) and (21 *b*) § 5·7, if we remember that W = − E, and that for n_1 and n_2 of § 5·7 we have now written n and j respectively.

It is of course possible to get the selection rules on the new theory also, but we shall not go into the details of this question here. They are the same as the rules (30 *a*) § 5 · 7 given by the old quantum theory.

11 · 8. *The Atom Under an External Magnetic Field. The Normal Zeeman Effect.*

In 1896 P. Zeeman discovered experimentally that in an external magnetic field of intensity H, a hydrogen line is split up into three components when viewed at right angles to the field, and two components when viewed along the field. If the frequency of the original line is ν_0, then the frequencies of the two displaced components are $\nu_0 + \Delta\nu$, $\nu_0 - \Delta\nu$, where

$$\Delta\nu = \frac{eH}{4\pi m_0 c}. \tag{1}$$

Immediately after Zeeman's discovery, and long before the advent of the quantum theory, H. A. Lorenz was able to account for this Zeeman effect by his classical electron theory. This is not surprising if we remark that the constant *h*, which is characteristic of the quantum theory, does not enter into the displaced frequencies (1). Bohr's quantum theory was equally successful in giving a solution of the problem. In 1916 Sommerfeld and Debye solved it by the method of the Hamilton-Jacobi equation effecting a separation of variables. A little later, Bohr gave a simpler theory with the help of the following theorem due to J. J. Larmor.

Larmor's Theorem.—The effect of a uniform magnetic field H on the motion of an electron under central forces (and some other forces also) is the same, to a first approximation, as if the magnetic field were absent and the whole system had a uniform rotation about an axis parallel to the direction of the field with the angular velocity $\dot{\psi}$ given by

$$\dot{\psi} = \frac{eH}{2m_0 c}.$$

Take the *z*-axis parallel to the direction of the field H. The components of the mechanical force on the electron at the points *x*, *y*, *z* due to the field are

$$\frac{eH}{c}\dot{y}, \quad -\frac{eH}{c}\dot{x}, \quad 0.$$

Let the components of the other force be X, Y, Z. The equations of motion are

$$m_0\ddot{x} = \frac{e}{c}\mathrm{H}\dot{y} + \mathrm{X}, \quad m_0\ddot{y} = -\frac{e}{c}\mathrm{H}\dot{x} + \mathrm{Y}, \quad m_0\ddot{z} = \mathrm{Z}.$$

Writing $\lambda = \dfrac{e\mathrm{H}}{m_0 c}$, and transposing, we get

$$m_0(\ddot{x} - \lambda\dot{y}) = \mathrm{X}, \quad m_0(\ddot{y} + \lambda\dot{x}) = \mathrm{Y}, \quad m_0\ddot{z} = \mathrm{Z}. \qquad (a)$$

Now suppose that the electron is free from the action of the magnetic field, but let the system rotate round the z-axis with a uniform angular velocity ψ, then the velocities of the electron are given by

$$u = \dot{x} - y\psi, \quad v = \dot{y} + x\psi, \quad w = \dot{z},$$

and the accelerations are given by $\dot{u} - v\psi$, $\dot{v} + u\psi$, $\dot{w}$, *i.e.*, by

$$\ddot{x} - 2\dot{y}\psi - x\psi^2, \quad \ddot{y} + 2\dot{x}\psi - y\psi^2, \quad \ddot{z}.$$

The equations of motion then become

$$m_0(\ddot{x} - 2\dot{y}\psi - x\psi^2) = \mathrm{X}, \quad m_0(y + 2\dot{x}\psi - y\psi^2) = \mathrm{Y}, \quad m_0\ddot{z} = \mathrm{Z}.$$

Suppose we take $\psi = \dfrac{\lambda}{2} = \dfrac{e\mathrm{H}}{2m_0c}$, then since c, the velocity of light, is very large, ψ is very small, and we can neglect ψ^2. The equations of motion then become

$$m_0(x - 2\dot{y}\psi) = \mathrm{X}, \quad m_0(\ddot{y} + 2\dot{x}\psi) = \mathrm{Y}, \quad m_0\ddot{z} = \mathrm{Z}. \qquad (b)$$

Since $2\psi = \lambda$, we see that the equations (a) are identical with the equations (b), and therefore Larmor's theorem is established.

The frequency ω_{L} of this " Larmor Precession " is then given by

$$\omega_{\mathrm{L}} = \frac{1}{2\pi}\cdot\frac{e\mathrm{H}}{2m_0c} = \frac{e\mathrm{H}}{4\pi m_0 c}. \qquad (2)$$

This additional frequency introduces a new quantum number m in the expression for the energy which explains the splitting up of the spectral lines.

We shall now give a detailed solution of the problem on the new mechanics, following the treatment given by Dirac.

We take a uniform magnetic field of intensity $\vec{\mathrm{H}}$, acting in the direction of the z-axis, so that

$$\mathrm{H}_x = 0, \quad \mathrm{H}_y = 0, \quad \mathrm{H}_z = \mathrm{H}, \quad \mathrm{E}_x = 0, \quad \mathrm{E}_y = 0, \quad \mathrm{E}_z = 0. \qquad (3)$$

Consequently, the scalar and vector potentials are

$$A_t = V = \frac{e}{r}, \quad A_x = -\frac{1}{2}Hy, \quad A_y = +\frac{1}{2}Hx, \quad A_z = 0. \tag{4}$$

In equation (9) § 11 · 6 we showed that the effect of the spin of the electron moving under an external electromagnetic field is to increase the potential energy by

$$\frac{1}{2m_0}\frac{h}{2\pi}\frac{e}{c}(\sigma_1 H_x + \sigma_2 H_y + \sigma_3 H_z) + \frac{1}{2m_0}\frac{ih}{2\pi}.\frac{e}{c}$$
$$\rho_1(\sigma_1 E_x + \sigma_2 E_y + \sigma_3 E_z),$$

which, in the present case, becomes, on account of (3) :

$$\frac{1}{2m_0}\cdot\frac{h}{2\pi}\frac{e}{c}\sigma_3 H. \tag{5}$$

Since the effect of the relativistic variation of the mass of the electron with its velocity is much smaller compared to the effect of the external magnetic field, we shall neglect the relativity effect, and take the classical Hamiltonian (17) of § 10 · 1 (4),

$$-eA_t + \frac{1}{2m_0}\left\{\left(p_x + \frac{e}{c}A_x\right)^2 + \left(p_y + \frac{e}{c}A_y\right)^2 + \left(p_z + \frac{e}{c}A_z\right)^2\right\}.$$

Adding the potential energy given by (5), and substituting the values from (4), we have for the Hamiltonian of the problem :

$$H' = \frac{1}{2m_0}\left\{\left(p_x - \frac{e}{2c}Hy\right)^2 + \left(p_y + \frac{e}{2c}Hx\right)^2 + p_z^2\right\}$$
$$-\frac{e^2}{r} + \frac{ehH}{4\pi m_0 c}\sigma_3. \tag{6}$$

If the magnetic field is not too large, the terms involving H^2 can be neglected, and (6) reduces to

$$H' = \frac{1}{2m_0}(p_x^2 + p_y^2 + p_z^2) - \frac{e^2}{r} + \frac{eH}{2m_0 c}(xp_y - yp_x) + \frac{eH}{2m_0 c}\frac{h}{2\pi}\sigma_3. \tag{7}$$

From (1) § 6 · 8 (1) we know that the component m_z of the orbital angular momentum is $m_z = xp_y - yp_x$. The Hamiltonian (7) can therefore be wirtten :

$$H' = \frac{1}{2m_0}(p_x^2 + p_y^2 + p_z^2) - \frac{e^2}{r} + \frac{eH}{2m_0 c}\left(m_z + \frac{h}{2\pi}\sigma_3\right). \tag{8}$$

The Hamiltonian of the hydrogen atom for no external field is simply $\frac{1}{2m_0}(p_x^2 + p_y^2 + p_z)^2 - \frac{e^2}{r}$, so that the extra terms due to the magnetic field are $\frac{eH}{2m_0c}\left(m_z + \frac{h}{2\pi}\sigma_3\right)$. Now in equations (7) and (8) of § 6 · 8 (1) it has been proved that each component of the orbital angular momentum, and therefore m_z, commutes with r and with $p_x^2 + p_y^2 + p_z^2$. σ_3, being a diagonal matrix, commutes, of course, with every matrix. The additional terms $\frac{eH}{2m_0c}\left(m_z + \frac{h}{2\pi}\sigma_3\right)$ commute, therefore, with the total Hamiltonian (8), and are thus constants of the motion, as shown in Chapter VI. Writing

$$H_0 = \frac{1}{2m_0}(p_x^2 + p_y^2 + p_z^2) - \frac{e^2}{r}, \tag{9}$$

and

$$H_1 = \frac{eH}{2m_0c}\left(m_z + \frac{h}{2\pi}\sigma_3\right), \tag{10}$$

the Hamiltonian (8) can be written

$$H' = H_0 + H_1, \tag{11}$$

where H_0 is the Hamiltonian for no external field, and H_1 is the additional term due to the field. As shown above, H_1 commutes with H', and is a constant of the motion. Therefore, the eigen-functions for the Hamiltonian H' will be those eigen-functions of H_0 which are also eigen-functions of H_1, since

$$H'\psi = H_0\psi + H_1\psi. \tag{12}$$

Moreover, if E, E_0, E_1 are the eigen-values corresponding H', H_0, H_1 respectively, then from (11) we have

$$E = E_0 + E_1, \tag{13}$$

showing that the presence of the magnetic field causes the stationary states to differ only in the value of the energy, and not in the form of the eigen-functions. This is the wave-mechanical analogue of Larmor's theorem.

Now the eigen-values of the orbital angular momentum m_z are given by the rule that the angular momentum is an integral multiple m of $\frac{h}{2\pi}$:

$$m_z = m\frac{h}{2\pi}, \quad (m \text{ integer}). \tag{14}$$

From the matrix for σ_3, its eigen-values are easily found to be $\sigma_3' = \pm 1$, because the eigen-values of a diagonal matrix are the terms in the leading diagonal.

Thus we find from (10):

$$E_1 = \frac{eH}{2m_0c}\left(m\frac{h}{2\pi} + \frac{h}{2\pi}\sigma_3'\right). \tag{15}$$

The selection rules for m_z now require that only those transitions are possible in which the magnetic quantum number m changes to $m + \Delta m$ such that

$$\Delta m = -1, 0, +1. \tag{16}$$

Since σ_3 commutes with every matrix, and also with the electric displacement, it will not change at all. We have therefore the result that in a transition E_1 changes to $E_1 + \Delta E_1$, where

$$\Delta E_1 = \frac{eH}{2m_0c}\cdot\frac{h}{2\pi}\Delta m. \tag{17}$$

Considering the Zeeman effect for a particular hydrogen line $\nu_0 = \frac{E_0}{h}$, we see on account of (13) and (17) that it is split up into the components $\nu_0 + \Delta\nu$, given by

$$\Delta\nu = \frac{\Delta E_1}{h} = \frac{eH}{4\pi m_0c}\Delta m, \tag{18}$$

where Δm has the three values (16). These displacements (18) are in agreement with experiments.

BIBLIOGRAPHY.

CHAPTER I.

THE THEORY OF RELATIVITY.

1. A. A. Michelson and E. W. Morley: *American Journal of Science* (3) **34** (1887), p. 333; *Phil. Mag.* (5) **24** (1887), p. 449.
2. H. A. Lorentz: *Amsterdam Proc.* **6** (1904), p. 809.
3. A. Einstein: *Ann. d. Physik.*, **17** (1905), p. 891.
4. Lorentz, Einstein, Minkowski: *The Principle of Relativity.*
5. A. Einstein: (1) *On the Special and General Theory of Relativity.*
   (2) *The Meaning of Relativity.*
6. A. S. Eddington: (1) *Space, Time and Gravitation.*
   (2) *The Mathematical Theory of Relativity.*
7. H. Weyl: *Space—Time—Matter.*
8. H. A. Lorentz: *Lectures on Theoretical Physics,* **3**.
9. M. von Laue: *Die Relativitatstheorie,* **1**.
10. Max Born: *Einstein's Relativity Theory and its Physical Foundations.*
11. R. C. Tolman: *Relativity, Thermodynamics and Cosmology,* Oxford (1934).
12. W. Pauli: *Enzyklopadie der math. Wissenschaften,* Bd. **V**. 2, p. 539, Art. 19.
13. R. Fürth: *Handbuch der Physik,* Bd. 12.

CHAPTER II.

GENERAL THEORY OF DYNAMICS.

1. A. Voss: *Enzyklopädie der math. Wissenschaften,* Bd. **IV**. 1, Art. 1 (Principles of Rational Mechanics).
2. L. Nordheim: *Handbuch der Physik,* Bd. **V**, Chap. 2.
3. L. Nordheim and E. Fues: *Ibid.*, Bd. **V**, Chap. 3 (Die Hamilton-Jacobische Theorie der Dynamik).
4. P. Appell: *Traité de mecanique rationelle,* Vol. 2.
5. E. T. Whittaker: *Analytical Dynamics.*
6. H. Poincaré: *Methode nouvelle de mecanique celeste,* Vol. 1.
7. E. Bloch: *L'ancienne et la nouvelle theorie des quaanta,* Chap. I (1930).
8. Max Born: *Mechanics of the Atom,* Chap. 1 (1927).

CHAPTER III.

STRUCTURE OF THE ATOM.

1. E. Rutherford: "The Structure of the Atom," *Phil. Mag.* (6), **21** (1911), p. 669; **27** (1914), p. 488.
2. C. Anderson: *Phys. Rev.,* **41** (1932), p. 405.
3. J. Chadwick: "Bakerian Lecture on the Neutron," *Proc. Roy. Soc.,* **142** (1933), p. 1.
4. W. Heisenberg: *Zs. f. Phys.,* **77** (1932), p. 1; **78** (1932), p. 156; **80** (1932), p. 587.
5. W. Bothe: "Das Neutron und das Positron," *Naturwissenschaften,* **21** (1933), p. 825.

6. W. Heisenberg: "Bemerkungen zur Theorie des Atomkerns," *Zeeman Verhandl* (1935), pp. 108–16.
7. E. Berthoud: *New Theories of Matter and the Atom.*
8. J. Kramers and H. Holst: *The Atom and the Bohr Theory of its Structure* (1923).
9. F. W. Aston: *Isotopes.*
10. G. Gamow: *Nuclear Physics.*
11. Max Born: *Atomic Physics* (1935).

CHAPTER IV.

THE QUANTUM THEORY OF RADIATION.

1. H. Herz: "Concerning an Effect of Ultra-violet Light on Electric Discharge," *Collected Works.*
2. Max Planck: Lecture to the German Physical Society, 14th Dec. 1900.
3. A. Einstein: "Zur Theorie der Lichterzeugung und Absorption," *Ann. d. Physik,* **17** (1905), p. 132; **20** (1906), p. 199.
4. Lord Rayleigh: *Phil. Mag.* (5), **49** (1900), p. 539.
5. J. H. Jeans: *Ibid.* (6), **10** (1905), p. 91.
6. A. H. Compton: *Phys. Rev.,* **21** (1923), p. 483; *Phil. Mag.* (6), **46** (1923), p. 897.
7. A. Sommerfeld: *Atomic Structure and Spectral Lines,* Chap. 1 (1935), 5th Ed.
8. A. H. Compton: "The Corpuscular Properties of Light," *Rev. Mod. Phys.,* **1** (1929), p. 74.
9. E. Fermi: "Quantum Theory of Radiation," *Ibid.,* **4** (1932), p. 87.
10. W. Heitler: "The Quantum Theory of Radiation" (1936).

CHAPTER V.

BOHR'S QUANTUM THEORY OF THE ATOM.

1. N. Bohr: (1) "On the Constitution of Atoms and Molecules," *Phil. Mag.* (6), **26** (1913), p. 1 and p. 476; (2) "On the Quantum Theory of Line Spectra," *Kopenhagen Acad.,* **4** (1918); (3) *The Theory of Spectra and Atomic Constitution, Cambridge* (1922); (4) *Three Lectures on Atomic Physics*; (5) *Ann. d. Physik,* **71** (1923), p. 228.
2. A. Sommerfeld: *Ann. d. Physik,* **51** (1916), p. 1; *Physikal. Zeits.,* **17** (1916), p. 491.
3. Max Planck: "Die Struktur des Phasenraumes," *Ann. d. Physik,* **50** (1916), p. 385.
4. P. Ehrenfest: *Physikal Zeits.,* **15** (1914), p. 657; *Ann. d. Physik,* **51** (1916), p. 327.
5. J. Stark: *Ann. d. Physik,* **43** (1914), p. 965.
6. P. Epstein: *Ibid.,* **50** (1916), p. 489; **51** (1916), p. 168.
7. W. Gerlach and O. Stern: *Zeits. f. Phys.,* **7** (1921), p. 249; **8** (1921), p. 110; **9** (1922), p. 349.
8. W. Pauli: "Über den Zusammenhang des Abschlusses der Electronen," *Ibid.,* **31** (1925), p. 765.
9. G. E. Uhlenbeck and S. Goudsmit: *Naturwissenschaften,* **47** (1925), p. 953; *Nature,* **117** (1926), p. 264.
10. A. Sommerfeld: *Atomic Structure and Spectral Lines,* Chaps. II, III (1935).
11. Max Born: *Mechanics of the Atom* (1927).
12. G. Birtwistle: *The Quantum Theory of the Atom,* Cambridge (1925).

13. F. Hund: *Limnienspektren und Periodisches System der Elemente,* Berlin (1927).
14. A. E. Ruark and H. C. Urey: *Atoms, Molecules, Quanta* (1930), Chap. III.
15. A. Smekal: "Allgemeine Grundlagen der Quanten—Statistik und Quantentheorie"; *Enzyklopädie der math. Wissenschaften,* Bd. **V**, 3, Art. 28 (1925).
16. W. Pauli: "Quantentheorie," *Handbuch der Physik,* Bd. **23** (1926).
17. A. Rubinowicz: "Ursprung und Entwicklung der Alteren Quantentheorie," *Ibid.,* Bd. **XXIV**, 1, 2nd Ed. (1933), Chap. 1.

CHAPTER VI.

HEISENBERG'S QUANTUM MECHANICS.

1. W. Heisenberg: *Zeits. f. Phys.,* **33** (1925), p. 879.
2. M. Born and P. Jordan: *Ibid.,* **34** (1925), p. 858.
3. M. Born, W. Heisenberg and P. Jordan: *Ibid.,* **35** (1925), p. 557.
4. P. A. M. Dirac: *Proc. Roy. Soc.,* **109** (1925), p. 642; **110** (1926), p. 501.
5. W. Pauli: *Zeits. f. Phys.,* **36** (1926), p. 336.
6. G. Birtwistle: *The New Quantum Mechanics,* Cambridge (1927).
7. M. Born and P. Jordan: *Elementare Quantenmechanik* (1930).
8. P. A. M. Dirac: *Principles of Quantum Mechanics* (1936), 2nd Ed.
9. W. Heisenberg: *Physical Principles of the Quantum Theory* (Mathematical Appendix) (1930).
10. A. E. Ruark and H. C. Urey: *Atoms, Molecules, Quanta* (1930).
11. E. V. Condon and P. M. Morse: *Quantum Mechanics* (1932).
12. J. von Neumann: *Mathematische Grundlagen der Quantenmechanik* (1932).
13. G. Temple: *Quantum Mechanics* (Methuen's Monographs).
14. P. Jordan: *Anschauliche Quantentheorie* (1936).
15. E. C. Kemble: "General Principles of Quantum Mechanics," *Rev. Mod. Phys.,* **1** (1929), p. 157; **2** (1930), p. 1.
16. E. Fues: "Einfuhrung in die Quantenmechanik," *Handbuch der Experimentalphysik,* Erganzungswerk, II (1935).

CHAPTER VII.

THE WAVE THEORY OF MATTER.

1. L. de Broglie: (1) "Ondes et Mouvement," *Thèses, Paris* (1924). (2) *Ann. d. Phys.,* **3** (1925), p. 22; (3) *J. de Phys.* (6), **7** (1926), p. 1.
2. E. Schrödinger: "Quantisierung als Eigenwertproblem," *Ann. d. Physik.,* **79** (1926), p. 484; etc.
3. C. Davisson and L. H. Germer: *Nature* (1927); *Phys. Rev.,* **30** (1927), p. 706.
4. G. P. Thomson: "Experiments on the Diffraction of Cathode Rays," *Proc. Roy. Soc.,* **117** (1928), p. 600.
5. O. Stern and I. Estermann, *Zeits. f. Phys.,* **61** (1930), p. 95; **73** (1931), p. 348.
6. E. Fues: "Beugungsversuche mit Materiewellen," *Handbuch der Exp. Physik* (1935), Erganzungswerk II.
7. L. de Broglie: *Introduction to Wave Mechanics.*
8. C. F. Meyer: "The Diffraction of Light, X-rays and Material Particles (1934).

CHAPTER VIII.

WAVE MECHANICS.

1. E. Schrödinger: *Ann. d. Physik* (4), **79** (1926), pp. 361, 489 and 734; Collected Papers on Wave Mechanics (1928).
2. A. Sommerfeld: *Wave Mechanics.*
3. P. A. M. Dirac: *Principles of Quantum Mechanics* (1936), 2nd Ed.
4. J. Frenkel: (1) *Wave Mechanics: Elementary Theory* (1930), (2) *Wave Mechanics: Advanced General Theory* (1933).
5. Arthur Haas: *Wave Mechanics.*
6. N. F. Mott: *Wave Mechanics.*
7. H. T. Flint: *Wave Mechanics* (Methuen's Monographs).
8. G. Castelfranchi: *Recent Advances in Atomic Physics,* vol. 2 (1932).
9. W. Pauliff "Die Allgemeinen Prinzipien der Wellenmechanik," *Handbuch der Physik,* Bd. **XXIV** (1933), Chap. 2, 2nd Ed.

CHAPTER IX.

PERTURBATION THEORY.

1. Born, Heisenberg, Jordan: *Zeits. f. Phys.,* **35** (1925), p. 557.
2. E. Schrödinger: *Ann. d. Physik.* (4), **80** (1926), p. 437.
3. P. A. M. Dirac: *Proc. Roy. Soc.,* **112** (1926), p. 661; **123** (1929), p. 714.
4. W. Heisenberg: *Zeits. f. Phys.,* **38** (1926), p. 411; **39** (1926), p. 499.
5. E. Fues: Störungsrechnung: *Handbuch der Physik,* Bd. **V.**
6. H. Weyl: *Gruppentheorie und Quantenmechanik* (1931), 2nd Ed.
7. A. Sommerfeld: *Wave Mechanics,* Chap 2, § 1-2.
8. P. A. M. Dirac: *The Principles of Quantum Mechanics,* Chaps. 8, 10.
9. J. Frenkel: *Wave Mechanics—Advanced General Theory,* Chaps. **V** and **VII.**
10. E. Wigner: *Gruppentheorie und ihre Anwendung auf die Quantumtheorie der Atomspektren* (1931).
11. Van der Waerden: *Gruppentheoretische Methode in der Quantenmechanik.*
12. C. Eckart: "The Application of Group Theory to the Quantum Dynamics of a Monoatomic System," *Rev. Mod. Phys.,* **2** (1930), p. 305.
13. H. Bethe: "Quantenmechanik der Ein-und Zwei-Elektronen Probleme," *Handbuch der Physik,* Bd. **XXIV,** 1 (1933), Chap. 3, 2nd Ed.

CHAPTER X.

PHYSICAL INTERPRETATIONS, AND HEISENBERG'S PRINCIPLE OF INDETERMINACY.

1. E. Schrödinger: "The Continuous Transition from Micro- to Macromechanics," *Naturwissenschaften,* **28** (1926), p. 664; *Ann. d. Physik.,* **82** (1927), p. 265.
2. Max Born: *Zeits. f. Phys.,* **37** (1926), p. 863; **38** (1926), p. 803; **40** (1927), p. 167.
3. W. Heisenberg: "Über den anschaulischen Inhalt der Quantentheoretischen Kinematik und Mechanik," *Zeits. f. Phys.,* **43** (1927), p. 172.
4. N. Bohr: *Naturwissenschaften,* **16** (1928), p. 245.
5. W. Heisenberg: *Physical Principles of the Quantum Theory* (1930).
6. Max Born: *Atomic Physics* (1935).
7. N. Bohr: *Quantum Theory and Description of Nature.*
8. Max Planck: *The Philosophy of Physics* (1935).

CHAPTER XI.

RELATIVISTIC QUANTUM MECHANICS.

1. P. A. M. Dirac: "Quantum Theory of the Electron," *Proc. Roy. Soc.*, **117** (1928), p. 610; **118** (1928), p. 351.
2. E. Schrödinger: *Ann. d. Physik*, **81** (1926); **82** (1927), p. 265.
3. W. Pauli: "Zur Quantenmechanik der Magnetischen Electrons," *Zeits. f. Phys.*, **43** (1927), p. 601.
4. W. Gordon: *Ibid.*, **40** (1926), p. 117; **48** (1928), p. 11.
5. C. G. Darwin: *Proc. Roy. Soc.*, **118** (1928), p. 654.
6. P. A. M. Dirac: *The Principles of Quantum Mechanics* (1936), Chap. **XII**, 2nd Ed.
7. A. Sommerfeld: *Wave Mechanics*, Chap. **II**, § 10.
8. G. Rumer: *Physikal. Zeits.*, **32** (1931), p. 601.
9. J. Frenkel: *Wave Mechanics*: *Advanced General Theory*, Chap. **VI**.
10. L. de Broglie: *L'electron Magnetique* (1935).
11. P. Jordan: *Anschaulische Quantentheorie*, Chap. **IV**.
12. A. S. Eddington: *The Relativity Theory of Protons and Electrons* (1936).

INDEX

(*Figures refer to page numbers*)

A

Action function 28
Action variables 41, 79
Adiabatic invariants 44, 79
Adiabatic principle 79, 107
Adjoint matrix 137
Adjoint operator 240
Advance of perihelion 86
Amplitudes 82
Anderson, C. 51, 264
Angle variables 41
Angular momentum 133, 268
Anomalous Zeeman Effect 99, 105
Antisymmetrical eigen-functions 214
Apse line 86
Aston, F. W. 50
Associated Legendre function 175
Associative law of matrices 113
Atomic magnetism 95
Atomic number 48
Atomic weight 48
Azimuthal quantum number 92

B

Balmer 72
Balmer lines 74, 204
Bechert 166
Blacket, P. M. S. 51
Black-body radiation 58
Bohr, N. 48
——— Quantum postulates 71
——— Theory of Hydrogen atom 72
——— Frequency relation 72, 73, 106, 124
——— Energy relation 72
——— Quantum condition 72, 106
——— Correspondence principle 80, 106
——— Magneton 97
——— Theory of harmonic oscillator 126
——— Complimentary character of waves and corpuscles 232
——— Zeeman Effect 278
Born, M. 105, 110, 115, 126, 227
Bose-Einstein statistics 217
Bothe and Geiger's experiment 54
Bothe and Becker 50
Burger, J. M. 44, 80

C

Canonical equations of motion 32, 35, 38, 117
Canonical transformations 36, 37, 117
Canonically conjugate variables 116, 124

C—(*Contd.*)

Causality 245
Central field of force 75
Chadwick, J. 50
Collision phenomena 105
Compton, A. H. 54, 63, 148
Compton Effect 54, 63, 148
Compton and Simon's experiment 54
Composition of velocities 11
Combination principle of Ritz 106, 108
Conservation of energy 26, 64, 65
Conservation of momentum 16, 64
Constancy of the velocity of light 6
Continuous spectrum 146
Continuum, four-dimensional 13
Contraction in length 3, 12
Core electrons 99
Corpuscular theory of light 53
Correspondence principle 80, 106
Crossed fields 105
Curl of a vector 250
Cyclic variables 41, 43

D

Darwin, C. G. 260, 270
Davisson, C. J. 157
Debye, P. 278
Debye-Scherrer apparatus 160
De Broglie, L. 133, 148, 156, 227
——— Conception of matter-waves 148
——— Physical interpretation 227
Deflection of α-particles 45
Deflection of electrons 236
Degenerate systems 44, 80, 191, 221
Determinism, classical 245
Diagonal matrix 114
Diffraction experiments 157
Diffuse series (spectra) 98
Differentiation of a matrix 115, 117, 122, 143
Dilatation in time 12
Dirac, P. A. M. 50, 51, 65, 100, 116, 120, 122
——— Quantum theory 117, 120, 122
——— Postulates 121, 122
——— Systems of similar particles 215
——— Statistics 217
——— Analysis of corpuscles and waves 232
——— Theory of the electron 261
——— Wave-equation 262, 263
——— Theory of the positron 264
——— Theory of the hydrogen atom 270

D—(*Contd.*)

Derac, P. A. M.
——— Theory of the Zeeman Effect 279
Dispersion phenomena 105
Distributive law of matrices 113
Divergence of a vector 250
Doppler Effect 237
Doubly ionised atom 74, 75
Duality between corpuscles, and waves 53, 54, 148

E

Eddington, A. S. 232
Ehrenfest, P. 79, 106
Eigen-functions 161, 189
Eigentime 14
Eigen-values 161, 189
——— of a matrix 114, 144
——— of an operator 240
Eigenvibrations 190
Einstein, A. 106, 133, 148, 227
——— relativity postulates 3
——— light quantum 54, 58
——— theory of photo-electric effect 54
——— statistics 217
Electron shells 102
Electric moment 81
Elliptic orbits 75, 84
Elsasser, W. 158
Energy levels 71
Epstein, P. 197
Equation of continuity 226, 250
Equations of motion—
Newtonian 16
Relativistic 19
Lagrangian 30
Hamilton's Canonical 32, 35
Quantum mechanical 121
Equivalence of observers 9
Euler's theorem 31
Events 13
Exchange degeneracy 211
Excited states of an atom 73
Exclusion principle (Pauli) 100, 217

F

Fermat's principle 154
Fermi, E.
——— statistics 217
——— element No. 93 103
Fine structure,
——— Sommerfeld's theory 83
——— Dirac's theory 270
Fine structure constant 90
Fitzgerald contraction 2
Four-dimensional vectors
——— position 21
——— velocity 22
——— momentum 22
——— force 23
——— electromagnetic potentials 251

F—(*Contd.*)

Frame of reference 6
Franck, J. 106
Frequency relation 72, 73, 106, 124

G

Galilean transformation 4
Gamma-ray microscope 234
Generalised co-ordinates 29, 119
Gerlach, W. 97
Germer, L. 159
Goudsmit 100
Gradient of a scalar 250
Ground orbit 73
Group velocity 153

H

Hamilton, W. R.
——— variation principle 25
——— Canonical equations 32, 42, 117
——— analogy between machanics and optics 154
Hamiltonian function 32, 35, 256
Hamiltonian operator 182
Hamiltonian-Jacobi equation 35
Harmonic Oscillator
——— Planck's 61
——— Bohr's theory 126
——— Heisenberg's theory 128
——— Schrödinger's theory 168, 170
Heisenberg, W.
——— theory of nuclear structure 52
——— critique of Bohr's theory 106, 148
——— postulates 106
——— multiplication rule 110
——— product 121
——— theory of harmonic oscillator 128
——— theory of Helium atom 209
——— perturbation theory 217
——— uncertainty relations 235
——— principle of indeterminacy 234
Heitler, W. 215
Hellinger, E. 145
Hermitian matrix 115, 138
Hermitian polynomial 172
Hertz, H—
Electro-magnetic waves 53, 251
Photo-electric effect 53, 61
Oscillating electric dipole 230
Hund, F. 215
Huygens' wave theory of light 53
Hydrogen atom—
Classical theory 69
Bohr's theory 79
Sommerfeld's theory 83
Schrödinger's theory 172
Dirac's theory 270
Hydrogen molecule ion 105
Hyperfine structure 52

I

Identity of mass and energy 21
Indeterminacy 234
Indicial equation 164
Inertia of energy 21
Infinite matrix 113, 126
Intensity of spectral lines 82, 230
Interval between events 13, 14
Invariance of Canonical equations 37
Invariance of wave-equation 10, 11, 264
Inverse matrix 112, 115
Inverse transformation 9, 112
Ionisation potential 69
Isotopes 48, 50

J

Jacobi-Hamilton differential equation 35, 36, 39
Jacobian determinant 117, 198
Jeans, J. H. 60
Jordan, P. 65, 110, 115, 126

K

K-shell 102
Kinetic energy—
  Newtonian 25
  Relativistic 20
Kirchhoff 59
Kunsman, C. H. 157

L

Lagrange's equations of motion 30
Lagrangian's function 27, 34, 256
Laguerre polynomial 175, 200
Lande, A. 98, 99
Larmor precision 278, 279
Laue's experiment 159
Leading diagonal of a matrix 114
Least action principle 28
Legendre polynomial 175
Liberation 42
Local time 14
Lorentz, H. A.—
  Contraction hypothesis 2
  Transformation 6
  Electron theory field equations 250
  Explanation of the Zeeman Effect 278

M

Magnetic moment of electron 95, 260, 264
Magnetic quantum number 99
Magneton (Bohr) 97
Mass-spectrograph 50
Mass-variation 16
Matrix 110
Matrix algebra 110, 137
Matrix analysis 140
Matrix function 140

M—(*Contd.*)

Maupertius' least action principle 29, 154
Maxwell's field equations 250
Mean value of a matrix 220
Mean value of an operator 242
Mendeleef's periodic table 48, 101
Michelson-Morley experiment 1
Micro-mechanics and macromanics 155
Millikan, R. A. 63
Minkowski, H. 13
Momentum—
  Conservation of 16
  Bohr's condition for angular 72
  Four-vector 22
  Orbital angular 133, 268
  Spin angular 260, 267

N

Negative energy states 264
Nernst, W. 132
Neutrinos 65
Neutron 50
Newton's corpuscular theory of light 53
Non-degenerate systems 43, 80, 190, 219
Normal states of an atom 73
Normalised wave-function 182
Nuclear motion 75
Nuclear structure 50, 51
Nucleus 46

O

Observables 106
Operators 179, 240
Orbits of electrons 106
Order of a matrix 112, 113
Orthogonal co-ordinates 40
Orthogonal eigen-functions 182
Ortho helium 215
Ortho terms 214

P

Parabolic co-ordinates 198
Para helium 215
Para terms 214
Paschen-Back Effect 99
Perihelion 86
Pauli, W. 50, 65, 99, 100, 101, 217, 260
Periodic systems 41
Periodic table of elements 101, 105
Peripheral electrons 48
Period of revolution 79
Permutation groups 215
Perturbation theory 188
Phase of vibration 56
Phase waves 148, 151
Photo-electric effect 53, 61, 62, 63, 148
Photo-electrons 61

P—(*Contd.*)

Photon 54, 55
Planck, Max 58, 71, 80
——— Quantum of action 54, 58, 61, 128
——— Quantum law 55
——— Constant 57, 61
——— Radiation law 60, 217
——— Determinism 246
Poincaré, H. 188
Poisson brackets 116
Polarisation of spectral lines 82
Polarisation rules 136, 231
Pole of a differential equation 163
Potential function 25
Potential, scalar 251
——— vector 251
Positron 51, 264
Principal quantum number 92
Principal series (spectra) 98
Pringsheim, P. 58
Probability 227, 228, 229
Proton 49

Q

Quadratic form (differential) 13
Quantum—
 of light 54
 of action 54
Quantum numbers 73
——— principal 92, 200
——— azimuthal 92
——— magnetic 99
——— spin 100
Quantum conditions 72, 76, 77, 80, 106, 121
Quotient of matrices 115

R

Radioactivity 48
Radioactive disintegration 65
Rayleigh's radiation law 60
Real operators 240
Reciprocity between observers 13
Relativistic equations of motion 19
Relativistic kinetic energy 20
Relativistic Lagrangian 34, 256
Relativistic Hamiltonian 35, 256
Relativistic transformation—
 for co-ordinates 6
 for velocities 11, 249
 for mass 16
 for forces 23
 for frequency 152
 for charge density 250
 for current density 250
 for electric intensity 254
 for magnetic intensity 254
 for electro-magnetic potentials 253
Relativity postulates 3
Resolving power of a microscope 235
Resonance 109, 209

R—(*Contd.*)

Resonance degeneracy 211
Rest energy 20
Rest mass of a particle 18
Ritz combination principle 74, 106, 108, 125
Rosette type of orbit 86
Rutherford, E. 45
——— scattering experiment 45
——— atom model 48, 71
Rydberg constant 72

S

Scalar potential 251
Scalar product of vectors 51
Scattering experiments,
——— α-particles 45
——— Compton 63
Schrodinger, E. 133, 148
——— Analogy between mechanics and optics 154
——— Wave-equation 150
——— Theory of harmonic oscillator 168
——— Theory of hydrogen atom 172
——— Perturbation theory 188
——— Theory of the Stark Effect 197
——— Physical interpretation 225
——— Wave packets 226
——— Relativistic wave equation 258
Schwarzschild, K. 80, 197
Secular equation 145, 223
Selection rules 83, 106, 136, 230
Self-adjoint equation 200
Series spectra 97
Shankland's experiment 65
Sharp series (spectra) 98
Simultaneous events 5
Singly ionised atoms 74
Singular points of differential equations 163
Skew-symmetrical matrix 115
Sommerfeld, A. 80, 98
——— Quantum conditions 80
——— Elliptic orbits 83
——— Theory of fine structure 83
——— Theory of conduction in metals 217
——— Theory of Zeeman Effect 278
Spin of the electron 100, 260
Spin quantum number 100
Spinors 261
Stark Effect 95, 133, 197
Stationary states 71, 106
Stephen's law 59
Stephen's constant 59
Stern, O. (diffraction of molecules 160
Stern-Gerlach experiment 97
Superposition principle 232
Superposition of waves 153, 232
Symmetrical eigen-functions 214
Symmetrical matrix 115
Symmetry character 216

T

Time dilatation 12
Time mean of a matrix 220
Third law of thermodynamics 132
Thomson, J. J. 45, 64
Thomson, G. P. 159
Threshold frequency 63
Toeplitz, O. 146
Transformation coefficients
  matrix 111
Transformation function 92
Transition 107
Transition probability 229

U

Uhlenbeck, E. 100
Uncertainty relations 235
Unit matrix 112, 114
Unitary matrix 139, 219

V

Van den Brock 48
Variation principle of Hamilton 29
Vector product of vectors 255
Vibrating membrane 189

W

Wave group 153
Wave-number 55, 73, 149
Wave-packet 226
Wave theory of matter 148
Wavicle 232
Wien's radiation law 60
Wigner, E. 215
Wilson Chamber 51

Z

Zeeman Effect 94, 278
——— anomalous 99, 105, 133
——— normal 133
Zero matrix 114

645--37 Printed at The Bangalore Press, Bangalore City,
by G. Srinivasa Rao, Superintendent.

To be had from :

THE TRANSLATION BUREAU
OSMANIA UNIVERSITY
HYDERABAD-DECCAN (INDIA)

Price Rs. 6 net
or sh. 8-6 net